A microeconomic model of banking with a strategically determined interbank market

Gideon Petrus du Rand

Dissertation presented for the degree of Doctor of Philosophy in Economics in the Faculty of Economics and Management Science at Stellenbosch University.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Supervisor: Prof Co-Pierre Georg

Co-supervisor: Prof Stanislaus Alexander du Plessis

Stellenbosch, South Africa

March 2020

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

This dissertation includes some portions of one unpublished paper. The development and writing of the unpublished paper were the principal responsibility of myself and, for each of the cases where this is not the case, a declaration is included in the dissertation indicating the nature and extent of the contributions of co-authors.

Declarations with regard to parts of the dissertation in which other authors were involved

Declaration by the candidate

With regard to chapter 2 (p. 33 – 34) and chapter 3 (p. 35 - 53), the nature and scope of my contribution were as follows:

Nature of contribution	**Extent of contribution**
Primary writing and analysis, with co-authors contributing general advice on direction of study and textual formulation suggestions	94%

The following co-authors have contributed to chapter 2 (p. 33 – 34) and chapter 3 (p. 35 - 53):

Name	**e-mail address**	**Nature of contribution**	**Extent of contribution**
Dr Toni Ahnert	toni.ahnert@gmail.com	General advice on direction of study and textual formulation suggestions	3%
Prof Co-Pierre Georg	cogeorg@gmail.com	General advice on direction of study and textual formulation suggestions	3%

With regard to section A.1, specifically on page 135, the nature and scope of my contribution were as follows:

Nature of contribution	**Extent of contribution**
Primary writing and analysis, with one co-author contributing a portion of the proof	95%

The following co-author has contributed to section A.1, specifically on page 135:

Name	**e-mail address**	**Nature of contribution**	**Extent of contribution**
Dr Toni Ahnert	toni.ahnert@gmail.com	A portion of the proof at the top of page 135	5%

Declaration with signature in possession of candidate and supervisor.

Declaration by the co-authors

The undersigned hereby confirm that

1. the declaration above accurately reflects the nature and extent of the contributions of the candidate and the co-authors to chapter 2 (p. 33 – 34), chapter 3 (p. 35 - 53) and section A.1 (p. 135),

2. no other authors contributed to chapter 2 (p. 33 – 34), chapter 3 (p. 35 - 53) and section A.1 (p. 135) besides those specified above, and

3. potential conflicts of interest have been revealed to all interested parties and that the necessary arrangements have been made to use the material in chapter 2 (p. 33 – 34), chapter 3 (p. 35 - 53) and section A.1 (p. 135) of this dissertation.

Signature	**Institutional Affiliation**	**Date**
Prof Co-Pierre Georg	University of Cape Town	26 September 2019
Dr Toni Ahnert	Bank of Canada	28 September 2019

Declaration with signatures in possession of candidate and supervisor.

Abstract

This dissertation generalizes the seminal model of financial contagion by Allen and Gale (2000) to allow an aggregate liquidity demand shock to occur with positive probability. A shock with positive probability can affect the ex ante portfolio choices of banks as well as the welfare of consumers. I numerically characterize the symmetric Nash equilibrium of the non-cooperative game between two representative regional banks. The solution fully characterizes banks' ex-ante optimal choices. I obtain the following results: (i) when the probability of the shock approaches zero, the allocation of Allen and Gale (2000) is obtained; (ii) in general, the equilibrium has three distinct characterizations, depending on the parameters: a no-default equilibrium, where no bank defaults; a single-default equilibrium where only the shocked bank defaults; and a mutual-default equilibrium, where the shock leads to contagion. When banks are able to internalize the ex-ante threat of a shock, contagion is rare: it is possible in, at most, 4% of the parameter space, and only for small shock probabilities.

Additionally, optimal risk-sharing is studied analytically in two novel aggregate benchmarks: a global bank with full information and a global bank with asymmetric information. A global bank with full information can observe consumer types. The allocation of a global bank with full information involves default after a large but sufficiently unlikely aggregate liquidity demand shock. Where default is not optimal, the allocation involves (i) holding excess liquidity when the shock is relatively likely, (ii) partial liquidation of the investment after a small and unlikely shock, and (iii) both excess liquidity and partial liquidation for shocks of intermediate size and probability. Under asymmetric information, a global bank cannot observe consumer types, and can offer less liquidity insurance than under full information.

Finally, when the numerically approximated Nash equilibrium is characterized by either no default or contagion, the decentralized solution attains the welfare of the benchmarks within numerical precision. However, when the Nash equilibrium is characterized by single default, the decentralized equilibrium is superior to the aggregate benchmarks. Thus, a global bank with regional branches can be inefficient for certain parameters in this model, relative to independent regional banks.

Opsomming

In hierdie proefskrif word die model van die oordraagbare verspreiding van finansiële probleme van Allen en Gale (2000) uitgebrei om 'n universele likiditeitskok met positiewe waarskynlikheid toe te laat. So 'n skok kan die ex ante keuse van bateportefeulje van banke asook die welsyn van verbruikers beïnvloed. Hierdie tesis los die simmetriese Nash ekwilibrium van die spel tussen twee verteenwoordigende plaaslike banke numeries op. Die belangrikste resultate is as volg: (i) die model lewer dieselfde verbruikerstoedeling as in Allen en Gale (2000) soos die waarskynlikheid van die algemene likiditeitskok na nul afneem; (ii) in die algemeen het die Nash ekwilibrium drie eienskappe, wat afhang van die parameters van die model: daar is 'n ekwilibrium waar geen bank bankrot gaan nie; 'n ekwilibrium waar slegs een bank bankrot gaan; en 'n algemene bankrotskap ekwilibrium, waar die bankrotskap van een bank oorgedra word en die bankrotskap van die ander bank veroorsaak. Selfs wanneer banke die ex ante risiko van 'n likiditeitskok kan antisipeer, is die ekwilibrium met oordraagbare bankrotskappe skaars: in slegs 4% of minder van die totale ruimte van modelparameters is oordraagbare bankrotskappe moontlik. Boonop gebeur oordraagbare bankrotskap slegs as die waarskynlikheid van 'n skok klein genoeg is.

Die optimale verdeling van risiko word ook bestudeer, deur middel van suiwer analitiese metodes. Twee nuwe welsynsmaatstawwe word op die makrovlak aangebied: 'n globale bank met volledige inligting en 'n globale bank met asimmetriese inligting. 'n Globale bank met volledige inligting kan die verbruikerstipe identifeer. As die skok groot en onwaarskynlik genoeg is, behels dié oplossing die bankrotskap van die globale bank, en derhalwe die volledige likwidasie van alle bates. Andersins het die oplossing een van die volgende eienskappe: (i) oor-

matige likiditeit, as die skok relatief onwaarskynlik is; (ii) gedeeltelike likwidasie van investering, as die skok klein en onwaarskynlik genoeg is, of (iii) beide oormatige likiditeit en gedeeltelike likwidasie van investering, as die skok gemiddeld groot en waarskynlik is. ’n Globale bank met asimmetriese inligting kan nie die verbruikerstipe identifiseer nie. Die resultaat is dat ’n globale bank met asimmetriese inligting minder likiditeitsversekering aan ’n verbruiker kan bied.

Laastens dui die resultate dat ’n enkele, globale bank in hierdie model sub-optimaal is wanneer die Nash ekwilibirium slegs een bankrotskap voorspel. Andersins lewer die Nash ekwilibrium die optimale uitkoms.

Acknowledgements

My thanks to Monique Reid, Dawie van Lill, Rochelle Jacobs, Estian Calitz, Andrie Schoombee and all my other colleagues who supported and encouraged me throughout the work on this dissertation. Thanks also to Peter Sinclair and all the participants at seminars at Stellenbosch University for valuable comments and suggestions.

Very special thanks to my supervisor and co-author Co-Pierre Georg, my co-supervisor Stan du Plessis and my co-author Toni Ahnert. Not only for the invaluable advice and direction they provided, but also for their cheerful attitudes that kept my spirits up during the most difficult parts of the work.

Lastly, thanks to my family for their love and support, and especially deep thanks to my partner Michael. This research has been as hard on him as it has been on me, but the support was unfailing and he was beyond patient with the many hours I had to spend at my desk. My deepest love and appreciation.

Contents

List of Figures

List of Tables

Introduction

An aggregate liquidity demand shock was at the heart of the recent global financial crisis. Specifically, there was an aggregate shock to the market liquidity of a number of bonds linked to subprime mortgages (Brunnermeier, 2009; Gorton and Metrick, 2012). Since these subprime bonds were widely used as collateral in bilateral repo transactions, the adverse shock to their market liquidity was amplified and spread across the financial sector (Brunnermeier and Pedersen, 2008; Copeland et al., 2014; Krishnamurthy et al., 2014; Martin et al., 2014). Even without infrequent catastrophic events such as financial crises, aggregate liquidity varies across the business cycle, as reflected in the time-varying spreads between liquid and illiquid assets (see, for example, Eisfeldt and Rampini, 2006). This dissertation contributes to the literature by studying the contagious spread of liquidity shocks across financial institutions, in particular, the theoretical literature focusing on deposit-taking banks that follows from the models of Diamond and Dybvig (1983), Allen and Gale (2000), Freixas et al. (2000) and Castiglionesi (2007).

The term *contagion* is commonly used in the financial intermediation literature to describe the spill-over of negative shocks in one part of the financial system to another, or to the real economy. Contagion is central to the arguments in this dissertation. However, the exact definition and measurement of contagion is not uncontroversial in the literature. In this dissertation, the definition of contagion corresponds with that of Forbes (2012), who documents the development of the usage and meaning of the term from 1990 onwards: Contagion is the spillover to financial intermediaries from "extreme negative events" that occurred elsewhere in the system. Common shocks or general imbalances are therefore ruled out[1]. Specifically, con-

[1]For instance, de Bandt and Hartmann (2000) argue for three different sources of systemic risk: (i) contagion, (ii) common shocks, and (iii) the general build up of widespread imbalances in the financial system.

tagion, as the term is used here, occurs when one financial intermediary defaults due to a large, localized shock, and this event induces the default of another intermediary (that did not experience the initial shock) that has positive balance sheet links to the aforementioned intermediary.

While the financial intermediation literature documents contagion across various types of financial intermediaries and markets, and resulting from various types of shocks, this dissertation focuses on deposit-taking banks. In the related literature, a deposit-taking bank is defined as a financial intermediary that offers an ex-ante welfare-improving demand-deposit contract to consumers who face privately observed idiosyncratic liquidity demand risk.

The Diamond and Dybvig (1983) model is the starting point of most studies of deposit-taking banks. In this three-period model there are consumers who face idiosyncratic liquidity demand risk; that is, they are initially uncertain about whether they will prefer to consume in the intermediate period (early consumers) or in the final period (late consumers). In the intermediate period, consumers privately learn their type (early or late). While the type of a specific consumer is private information, the aggregate proportions of early and late consumers are common knowledge. Therefore, there is no aggregate uncertainty. There is a single deposit-taking bank that can pool the resources of consumers, and can offer them the first-best risk-sharing allocation, which improves on the outcome that consumers can obtain via private trade. The mechanism the bank employs is a demand-deposit contract, where all withdrawals in the intermediate period receive the same return, unless the bank fails. In other words, the deposit returns are risk-free (equivalently, non-state-contingent), providing that the bank survives to honour the contract.

Risk-free deposit returns is a key assumption in this dissertation and in much of the contagion literature[2]. This assumption serves as a reduced-form way of capturing the feature of real world demand-deposit contracts that offer a known, risk-free return that is independent of the moment of withdrawal, conditional on a bank not going into default. Such a non-contingent demand-deposit contract can also be motivated as an optimal mechanism. For instance, Diamond and Rajan (2001) show that non-contingent demand-deposits arise as a natural commit-

[2] E.g. Diamond and Dybvig (1983) Allen and Gale (1998, 2000, 2004*a*,*b*), Freixas et al. (2000), Dasgupta (2004), Castiglionesi (2007), and Allen et al. (2009).

ment mechanism to prevent banks from extracting surplus after receiving deposits.

The assumption of a non-contingent demand-deposit also provides a mechanism for modelling the causes of bank failure. In Diamond and Dybvig (1983), the non-contingent demand-deposit contract is augmented with a sequential-service constraint: consumers who attempt to withdraw in the intermediate period are served in the (random) order in which they lodge their claims. This induces multiple equilibria, one in which there is a run on the bank, where all consumers attempt to withdraw early due to self-fulfilling beliefs. In other models[3], as in the one studied in this dissertation, it is assumed that there is uncertainty about the proportion of consumers who wish to withdraw early. Whatever the mechanism, when the early demand is so large that a bank has insufficient resources to offer all consumers at least the non-contingent deposit return, a run occurs and the bank defaults.

The contagion literature typically adds two components to the basic Diamond and Dybvig (1983) setting: (i) more financial institutions, and (ii) risk at the aggregate as well as the institutional level. Institution-level risk creates an incentive for the financial institutions to form balance sheet links in order to cross-insure against individual risk, while the aggregate risk can induce contagion. A common feature in the theoretical literature on contagion is the *resilient-yet-fragile* property: balance sheet linkages provide a buffer against small shocks, but provide the mechanism for the contagious spread of larger shocks[4]. Therefore, the implications of direct balance sheet links between banks for financial fragility are ambiguous: they depend on the reason for and size of the links. The default at a specific bank occurs if the initial shock is large enough relative to the ex ante buffers held by that bank, but contagion (or a banking crisis) ensues if the balance sheet links across banks are large and pervasive enough to cause other banks to fail. Therefore, the ex-ante portfolio choices of banks, which include their balance sheet links to other banks, are paramount to understanding the risk of contagion.

The model in this dissertation allows the study of the ex-ante portfolio choices of banks that may lead to contagion. The starting point is the seminal models of financial contagion of

[3]E.g. Allen and Gale (2000),Dasgupta (2004),Ennis and Keister (2006) and Castiglionesi (2007).

[4]See, for example Gai and Kapadia (2010), Elliott et al. (2014), Acemoglu et al. (2015) (and the extensive literature that follows) who study contagion in more realistic network settings.

Freixas et al. (2000) and in particular of Allen and Gale (2000)[5]. The economy extends over three dates. There are consumers with preferences as in Diamond and Dybvig (1983), and all consumers are initially identical. In the intermediate period, consumers learn their types as private information: a proportion of them wishes to consume at the interim date (early consumers) and the remainder wishes to consume at the final date (late consumers). The consumption preferences of consumers are initially unknown, and banks provide insurance against such idiosyncratic liquidity risk by offering demand-deposit contracts.

The bank chooses its portfolio at the initial date, allocating its resources between liquidity, a productive asset and an interbank deposit. Liquidity is simply a storage technology with zero net return in all periods. The productive investment has a positive net return if held to maturity (until the final date) but a negative net return if liquidated early (at the interim date). The interbank deposit has an endogenous return that depends on the equilibrium choices of the banks and the state of nature that realizes.

There are two regions, and in the absence of an aggregate liquidity demand shock, some regional variation in liquidity demand, which motivates insurance across regions. When an aggregate liquidity demand shock occurs, by contrast, there is a higher aggregate proportion of consumers who wish to consume at the interim date. The aggregate shock cannot be diversified away across regions.

The model in this dissertation follows that of Allen and Gale (2000) as closely as possible, but contributes by augmenting their model in one important aspect: it introduces aggregate liquidity demand shocks with *positive* probability. This is not a trivial addition to the model, because in their model, the aggregate liquidity demand shock is completely unanticipated (it has zero probability). This means that the shock can have no impact on the ex-ante choices of banks or the ex-ante welfare of consumers. Furthermore, banks attain the deterministic first best allocation, so that there can be no ex-ante role for policy. When the aggregate liquidity demand shock occurs with positive probability, banks can internalize the risk, which affects their

[5]Both models yield very similar results and were published at the same time. The focus is on the model of Allen and Gale (2000) as the benchmark because it yields more tractable analytical solutions. The methodology, however, could just as well be applied to the equally appealing model of Freixas et al. (2000).

ex-ante portfolio choices. This generalization of the Allen and Gale (2000) model therefore allows the study of the impact of simultaneous variation in the probability and size of aggregate liquidity risk on portfolio choices and on risk-sharing across states and over time. It also allows for a possible role for policy to improve the outcome over that which banks would choose independently.

The analysis of Allen and Gale (2000) shows that, in the absence of aggregate liquidity demand shocks, the first-best allocation is deterministic (both early and late consumers face no consumption risk) and incentive compatible (where the consumption of late consumers is at least as large as that of early consumers, so that late consumers have no incentive to run on the bank). The first-best allocation can be feasibly decentralized if banks hold interbank deposits in each other that are large enough to provide the necessary liquidity transfers to cover regional liquidity demand shocks. Next, they show that, if an unanticipated and sufficiently large aggregate liquidity demand shock hits a representative bank in one region, it will be unable to offer an incentive-compatible allocation. This induces a run on that bank, so that the bank defaults and liquidates fully. Lastly, depending on the structure of the network of cross-holdings of interbank deposits, a default in one region can spill over to cause a default in the next region that did not experience the aggregate liquidity demand shock. In Allen and Gale (2000), contagion occurs only when the network is sufficiently incomplete.

The key contribution of this dissertation is the full characterization of strategic interbank market determination and the prevalence of contagion in a situation where the probability of the aggregate liquidity demand shock is positive. This is novel in the literature. Other studies of interbank market contagion assume no interbank deposit market[6], or a perfectly competitive interbank market[7], or they study a different solution concept[8].

The solution concept of the model studied here is the symmetric Nash equilibrium of the non-cooperative game between two representative regional banks[9] subject to regional and

[6]E.g. Diamond and Rajan (2011), Duarte and Eisenbach (2018) and Acharya and Yorulmazer (2008). In these models banks are exposed to each other via other channels than direct balance sheet links, such as correlated investment, fire sales of liquid assets or correlated information.

[7]E.g. Allen and Gale (1998) and Allen et al. (2009)

[8]E.g. Dasgupta (2004) who studies a global games coordination equilibrium.

[9]This is a simplified version of the model in Allen and Gale (2000) where there are four regions, each with a

aggregate liquidity risk. The two regional banks independently choose their asset portfolios (holdings of liquidity, productive investment, and the interbank deposit in the demand-deposit contract of the bank in the other region), as well as the deposit return on early withdrawals (which is subject to a non-state-contingency constraint unless default occurs). As a bank cannot observe the type of consumer, it must offer an incentive compatible contract in any state where it wants to avoid a run. That is, late consumers must receive a greater return by waiting until the final period to withdraw. Otherwise they will attempt to withdraw in the intermediate period, which constitutes a run on the bank.

Ex ante, there are three distinct types of symmetric Nash equilibria that may arise: one in which neither bank ever defaults, one in which only the bank that experiences the aggregate liquidity demand shock defaults, and an equilibrium where both banks default whenever either bank experiences the aggregate liquidity demand shock. The last situation represents the definition of contagion in this model. As in Allen and Gale (2000), the aggregate liquidity demand shock is localized in one region. In the contagion case, the bank that experiences the aggregate shock does not hold enough resources to offer an incentive-compatible contract that satisfies the non-state-contingency constraint on deposit returns. It thus experiences a run, and defaults. The consequent return on the interbank deposit of the other bank (that did not experience the aggregate liquidity demand shock) is so much lower that it also cannot offer an incentive-compatible return on the deposit contract, and thus it also faces a run and defaults. The contagion case is thus equivalent to the mutual-default equilibrium type in this dissertation.

The formal problem facing a representative bank in an arbitrary region is ex-ante fully symmetric across the two regions. This defines a single symmetric best-response function that is applicable to the representative bank in either region. The symmetric Nash equilibrium of the game is defined by the fixed point of this symmetric best-response function.

This problem is extremely difficult to study analytically due to the structure of the best-response function. The best-response function is complicated as a result of several potentially

representative bank. The difference to the setup is immaterial for the results obtained, but studying a two-player game is analytically and computationally more tractable than studying a four-player game.

binding constraints faced by the banks. The full analytical characterization of the fixed point of the best-response function may require up to 36 distinct cases, based on the different patterns in which the many constraints may bind or be slack. Hence, the symmetric Nash equilibrium of the full strategic game between the two banks is approximated numerically across the full range of model parameters.

The numerical approximation of the symmetric Nash equilibrium is an iteratively stable fixed point of the implied symmetric best-response function of one bank to the choice vector of the other bank. The best-response choice of a bank (to a given choice of the other bank) is constructed by direct numerical optimization of the constrained objective function of an arbitrary bank. The constrained objective function is given by the strategic problem of an ex ante symmetric bank in an arbitrary region. The constrained objective function of a regional bank is the regional ex-ante expected utility of an arbitrary consumer in that region, subject to a set of incentive-compatibility and non-negativity constraints.

For each parameter set and for each of the three potential equilibrium types independently, standard constrained numerical optimization routines are used to solve for the approximate best response of the representative bank in one region to a feasible initial choice of the bank in the other region. The algorithm then replaces the choice of the counterparty bank with the found best response and then iterates over the best-response function until the sequence of best responses converges to a fixed point, within numerical precision. This yields a candidate symmetric Nash equilibrium for each potential equilibrium type. Finally, the approximate symmetric Nash equilibrium type is selected from the three candidate types as the one that attains the maximum expected utility for the given parameter set.

The numerical approach yields several results in the decentralized setting:

(i) As the probability of aggregate risk approaches zero, the numerically approximated symmetric Nash equilibrium converges on the deterministic first-best allocation, replicating the analytic results of Allen and Gale (2000). This serves as an important validation of the reliability of the numerical approach.

(ii) There are parameter ranges in which contagion never occurs in equilibrium, given the positive probability of the aggregate liquidity demand shock. Specifically, this occurs when the aggregate liquidity demand shock is small enough.

(iii) If the aggregate liquidity demand shock is large enough to make contagion possible, the symmetric Nash equilibrium has one of three different characterizations[10]: if the probability of the shock is small enough, the equilibrium is characterized by both banks defaulting if the aggregate shock realizes (i.e. contagion occurs). If the probability is intermediate, the equilibrium is characterized by only one bank (the one hit by the aggregate liquidity demand shock) defaulting (i.e. default without contagion occurs). If the probability of the aggregate liquidity demand shock is high enough, the banks choose portfolios and deposit returns such that neither bank ever defaults.

(iv) Across parameter ranges where the three equilibrium types have positive measure, optimal choices are discontinuous functions of the parameterization, i.e. at the parameter boundary where one equilibrium type switches to another, there is a discontinuity in the optimal choice. This mirrors similar discontinuities in the aggregate benchmark allocations where the solution switches from one type to another (the benchmarks are discussed below).

(v) The endogenous interbank deposits are always symmetric and positive, which implies that the interbank network is always complete in equilibrium. This contributes to the study of contagion in the endogenous network formation literature[11]. Thus, when contagion occurs, it occurs in a complete interbank network, in contrast to Allen and Gale (2000).

(vi) Contagion is a rare phenomenon in the parameter space of this model. In an exercise where model parameters where drawn independently from their allowable ranges, contagion is possible in approximately 4% of the parameter draws. This is an upper bound on the prevalence of contagion, as one of the randomly drawn parameters is the proba-

[10] These different equilibrium types have positive measure in the numerically studied parameter space.

[11] E.g. Jackson and Wolinsky (1996), Bala and Goyal (2000), Dutta and Jackson (2003), and Bloch and Jackson (2006).

bility of the aggregate liquidity demand shock. Contagion only occurs in a parameter set where it is possible *if* the aggregate liquidity demand shock is realized. Taking into account the probability of the shock in parameter sets where contagion is possible, the ex ante likelihood of contagion falls to approximately 0.5%.

To analyze the welfare outcomes in the numerically approximated decentralized equilibrium, a novel aggregate benchmark is presented as a comparison allocation. This allocation is fully characterized analytically, and is augmented with numerical examples to provide a full comparison of the outcomes in the benchmarks with those from the numerically approximated decentralized equilibrium. The benchmark is labelled a global bank and is studied under two information assumptions: under full information and under asymmetric information.

A global bank is defined as single banking company with two branches, one in each region[12]. Therefore, the two regional banks in the decentralized case become the regional *branches* of a global bank in the aggregate benchmark. The objective of a global bank is to maximize the ex ante welfare of a consumer from an arbitrary region. As such, a global bank invests in its portfolio at the aggregate level, and treats consumers from different regions identically[13]. Importantly, this means a global bank either survives as a whole, or fails as a whole. It cannot allow default in just one region. In the decentralized situation, in contrast, it is possible for the bank in one region to fail while the bank in the other region survives.

I study the global bank benchmark under two information assumptions: under full information, a global bank is not subject to the information constraint: it can observe the type of consumer (early or late). This means it can force a late consumer to withdraw in the final period and hence, avoid a run by assumption. The implication of this is that a fully informed global bank is not subject to an incentive compatibility constraint - it can offer late consumers

[12]This benchmark can loosely be related to the empirical work on the historical features of national banks by Calomiris and Carlson (2016, 2017). While national banks are somewhat less relevant in economies like the US, they are still important in countries with more concentrated banking systems like South Africa.

[13]The global bank allocation is therefore different from a standard *benevolent global bank* allocation that is typically used as an aggregate benchmark. Such a global bank allocation would be be subject only to the defining characteristics of the economy studied. It would be able to treat consumers in different regions differently and offer state contingent consumption allocations to all consumers. This benchmark is formally identical to the autarkic allocation characterized in Castiglionesi et al. (2017), and hence omitted from the dissertation.

a lower return than early consumers without experiencing a run and defaulting. However, a fully informed global bank is still subject to offering a non-state-contingent early consumption *unless it defaults* (default of a global bank is the aggregate benchmark equivalent of contagion in the decentralized setting).

One of the contributions of the dissertation is to show that the non-state-contingency constraint on early consumption is sufficient to make default optimal for a global bank with full information (for some parameter regions). Thus, even though a fully informed global bank can always avoid a run (and default), it is sometimes optimal for a global bank to *choose* to allow a run and default. This is because default is the only tool available to the global bank whereby consumption risk can be transferred from late to early consumers. The optimality of default in this dissertation is similar to results on optimal default and financial crisis in Allen and Gale (1998).

The global bank allocation under full information represents the optimal risk-sharing arrangement when the types of consumers are observed, but with the constraint that early consumption is non-state-contingent, except if the global bank optimally chooses to default in the intermediate period. The ex-ante choice of whether or not to default in the intermediate period thus is a choice between two *regimes* that are distinct because each is subject to a different set of economically motivated constraints. In the no-default regime, the global bank is subject to the constraint of non-state-contingent early consumption; in the default regime, it is not. The solution to the global bank problem can thus have one of two distinct possible characterizations, depending on whether the default or the no-default regime is ex-ante optimal. Moreover, at the boundary of the parameter regions where the optimal regime switches from one type to the other, the optimal choices of a global bank can be discontinuous.

For an aggregate liquidity demand shock that is both sufficiently large and sufficiently unlikely, a global bank optimally chooses to default and fully liquidate all assets in the event of such a shock (under both full and asymmetric information). All proceeds from the liquidation of asset is then proportionally distributed to all consumers. Effectively, this is the only tool that allows a global bank to avoid the constraint that a deposit contract must provide risk-free deposit returns at the interim date, which can be quite costly ex ante. This result agrees with the insights of Allen and Gale (1998), who emphasize the value of bank default as a tool to manage aggregate solvency shocks.

In contrast, when a global bank optimally does not choose to default after an aggregate liquidity demand shock, there are two tools that balance the marginal utility of early and late consumers. These tools are labelled *excess liquidity* and *partial liquidation*. Excess liquidity may be optimally used in the state in which the aggregate liquidity demand shock does not realize. In this case, some liquidity is used to finance late consumption (which is ex-post inefficient). In contrast, partial liquidation of the productive investment may be necessary after an aggregate liquidity demand shock does realize. In this case, some of the productive investment is liquidated early in order to finance early consumption (which is also ex-post inefficient). The extent of the usage of these two ex-post inefficient tools depend on the probability of an aggregate liquidity demand shock, which in turn determines their relative ex-ante efficiency. For a likely aggregate liquidity demand shock, a global bank holds only excess liquidity. For a small and unlikely shock, however, only partial liquidation is used. For shocks of intermediate size and probability, a global bank uses both excess liquidity and partial liquidation. This is different from the study of Allen and Gale (2000), where liquidity is used exclusively to finance early consumption, and the productive investment is used exclusively to finance late consumption.

Next, I study the global bank under asymmetric information. A global bank that cannot observe consumer types must offer an incentive compatible consumption allocation to avoid a run. In other words, late consumers must receive a larger consumption allocation when they withdraw in the final period than early consumers receive by withdrawing in the intermediate period, otherwise all consumers would attempt to withdraw in the intermediate period and the bank will default. I show that this information constraint is economically important by characterizing the parameter regions where it binds and analyzing how the choices of a global bank under asymmetric information differs from its choices under full information. The results show that, due to the incentive compatibility constraints, a global bank that cannot observe consumer types relies more on excess liquidity and less on partial liquidation, both on the intensive margin (levels) and on the extensive margin (for a larger range of parameters). As a result, there are parameter regions where a global bank can provide less insurance against liquidity risk to consumers under asymmetric information than it can under full information.

The final contribution of this dissertation is a comparison of the welfare outcomes in the analytical aggregate benchmarks with the numerical results in the decentralized problem. There are two key findings:

(i) When the parameter set is such that the Nash equilibrium of the decentralized problem is characterized by either no default or mutual default (i.e. contagion), there is no evidence that the decentralized allocation is inferior to the benchmark. Therefore, for these parameter sets, the decentralized equilibrium attains the same level of welfare as the global bank benchmark.

(ii) When the parameter set is such that the Nash equilibrium of the decentralized problem is characterized by single default, the decentralized allocation is often significantly superior to the global bank allocation (under full or asymmetric information). Since the global bank is a single bank with two regional branches where consumers must be treated identically, it loses a degree of freedom; that is, the global bank does not have the option of allowing the branch in one region to fail while the other survives. This extra degree of freedom in the decentralized problem yields the superior outcomes. This result is tangentially related to results in the literature where banks can become inefficiently similar either by excessive integration on the interbank market, which exposes all banks to the same aggregate risk (Castiglionesi et al., 2017); or, by excessive diversification in investment, so that all banks face the same returns on investments and their depositors receive identical returns (Wagner, 2010).

This dissertation is organized as follows. Chapter 1 presents a review of the related literature. It reviews in some depth the closest papers to the model presented here. These are models that have considered banking and contagion in the face of either pure liquidity risk or a combination of liquidity and investment risk. The following sections of the literature review covers other possible sources of contagion: correlated investments, fire sales of assets, and correlated information. A final section reviews the literature on the impact of interconnectedness on contagion. Chapter 2 sets out the environment that defines the economy subject to liquidity risk, and chapter 3 provides the analytical results on the benchmarks in this economy. Chapter 4 sets out the strategic problem of independent regional banks, provides a discussion the nature of the interbank market and defines the Nash equilibrium. It concludes by presenting a discussion on the numerical approach used to solve for the Nash equilibrium. Chapter 5 presents the results of the characterization of the numerically approximate symmetric Nash equilibrium of the strategic interbank problem. Chapter 6 provides a comparison of the welfare outcomes of the decentralized Nash equilibrium with those of the aggregate benchmarks, and a final chapter provides a conclusion.

Chapter 1

Literature Review

This literature review is structured as follows. In the first section, I provide a review of historical banking crises, including the global financial crisis of 2007/8, and the central role of liquidity uncertainty in the propagation of localized shocks that result in large-scale financial contagion. To study this effect, economists have proposed numerous models based on pure liquidity risk, or a combination of liquidity and investment risk. The second section provides a review of models taken from the strand of literature that is closest to the work done in this dissertation, which consider only pure liquidity risk. The third section considers models that are also close to the one in this dissertation, except that they consider a combination of liquidity and investment risk. The fourth section considers models where financial contagion can arise for reasons other than pure liquidity or investment risk. These are situations where contagion can arise due to (i) correlated investments, where banks invest in similar (illiquid) types of projects, (ii) fire sales of assets, where banks structure the asset side of their balance sheets using similar assets that are subject to price disturbances in crises, and (iii) correlated information, where a shock to one part of a banking system reveals information about other parts of the system. In the final part of the review I shift attention away from shocks to the banking system to consider the interconnectedness of the financial sector as a mechanism for the propagation and augmentation of localized shocks.

1.1 The global financial crisis and why liquidity matters

Brunnermeier (2009) places a "liquidity crunch" at the centre of the global financial crisis of 2007/8. He identifies the vicious liquidity spirals that occurred during the crisis: as falling asset prices reduced the capital positions of financial institutions, this led to liquidity hoarding and reductions in lending. Combined with a high degree of interconnectedness, these liquidity shocks spread contagiously throughout the financial sector, causing massive reductions in the stock market value of financial institutions, further reducing lending, inducing real reductions in output and consumption, and increases in unemployment.

The primary vector of contagion during the crisis was the broad reliance of the interconnected banking sector on short term financing via the collateralized repo market (Copeland et al., 2014; Krishnamurthy et al., 2014; Martin et al., 2014). Distress in the mortgage market caused panic over the value of securitized mortgage backed securities, which served as collateral for a substantial proportion of the repo market. For instance, Gorton and Metrick (2012) present evidence of a "run on repo", documenting a massive increase in the margins in repo transactions which corresponds to a dry up of interbank liquidity during the global financial crisis – the margins became so high that the assets backing a potential repo transaction had essentially no value as collateral. Banks therefore could not use a large portion of their assets as collateral in short term borrowing, and hence could not generate the liquidity needed to satisfy short term liquidity demand. Thus, a localized shock in the housing sector led to large scale liquidity shortages in the banking sector.

The reliance of the banking sector on the repo market is a relatively recent phenomenon, but bank crises and financial contagion are not (Friedman and Schwartz, 2008; Minoiu and Reyes, 2013). Direct interbank lending has also been a source of interconnectedness, which has led to contagion in historical banking crises. Similar to the global financial crisis, aggregate liquidity reductions that spread across banks connected via direct interbank deposits occurred during the Great Depression. Mitchener and Richardson (2013, 2019) document how liquidity shocks to peripheral banks led to a "cascade of interbank withdrawals" (and therefore a reduction of aggregate liquidity in the system) that spread first to the regional financial centres and then to central reserve cities.

Several theoretical reasons for aggregate liquidity problems have been suggested. I review two salient contributions that argue for two distinct reasons why banks might hoard liquidity and hence cause aggregate liquidity disturbances: Acharya and Skeie (2011) present the argument that, during a crisis, banks might hold on to liquidity rather than lend it to counter party banks that are in need of liquidity for precautionary reasons, regardless of the level of counter party risk, whereas Heider et al. (2015) argue that banks hoard liquidity because they fear an increase in counterparty risk during a crisis and not for precautionary reasons.

Acharya and Skeie (2011) show theoretically that shocks to asset valuations can cause liquidity hoarding, or even a complete market freeze, with correspondingly large interbank interest rate premiums. These effects require that banks with high leverage face moral hazard problems - the incentive to hoard is induced by the risk (after the shock) that the interbank loans of highly leveraged banks may not be rolled over after the shock, therefore such banks keep liquidity for precautionary reasons rather than lending to other banks[1].

Another mechanism that may lead to aggregate liquidity problems is an increase in counterparty risk during a crisis. Heider et al. (2015) present a model where liquidity hoarding occurs due to increased counterparty risk stemming from an adverse selection problem, rather than for precautionary reasons. In their model, the level of aggregate liquidity available to banks is an endogenous consequence of their decisions. The adverse selection problem arises due to bank specific investment risk that is only privately observable (to the investing bank). When interest rates are high, only riskier banks want to borrow which could lead to reduced liquidity available on the interbank market, or even to the complete collapse of the interbank market[2].

It is an empirical question which of these two mechanisms – precautionary liquidity hoarding or liquidity hoarding due to increased counterparty risk – better explain the aggregate liquidity problems of the global financial crisis. Afonso et al. (2011) present an empirical study

[1] See also Allen et al. (2009), Caballero and Krishnamurthy (2008) and Diamond and Rajan (2011)

[2] See also Flannery (1996) and Freixas and Jorge (2008).

based on the daily transactions of banks on the Federal Funds Market. They found that the main reason for the reduction in liquidity was counterparty risk, as riskier banks did not hoard liquidity.

From the literature reviewed here, two key ideas are central to financial contagion based on the most recent as well as historical banking crises[3]: financial contagion between banks has been associated with aggregate liquidity disruptions, and the degree and cause of interconnectedness materially affect the severity of the ensuing crisis.

This dissertation contributes to the theoretical literature on financial contagion that assumes exogenous aggregate liquidity uncertainty. The main contributions are to provide a framework for understanding the endogenous formation of interconnections conditional on aggregate liquidity risk, and to study the consequences for contagion of these endogenous links. Since the dissertation contributes to the theory of financial contagion, this review turns now to the precursors of the model described in later chapters.

1.2 Models with only liquidity risk

This dissertation adds to a narrow literature that considers contagion to be a consequence of liquidity disturbances, based on the seminal theoretical structure of Diamond and Dybvig (1983) regarding banks that provide liquidity insurance.

The closest models to the one considered in this dissertation are those in Allen and Gale (2000), Dasgupta (2004) and Castiglionesi (2007). These models all build on the standard three period setting of Diamond and Dybvig (1983), where information revealed in the intermediate period yields the outcome of either an individual bank run or contagion. In all the papers reviewed in this section, there are two assets: a risk-free storage technology (called *liquidity* in this dissertation) and a productive investment that can yield a higher return in the final period

[3]Of course, the global financial crisis and all previous banking crises are far more multifaceted than the brief treatment devoted to them here, but the focus in this dissertation is not on such issues as risk-shifting, securitization or political influence on the mortgage market. For reviews of these issues see, for example, United States Financial Crisis Inquiry Commission (2011).

than liquidity. Various papers make different assumptions on the productive investment technology: it may be risky or risk free, liquid or illiquid. I provide the details for each case below.

A central assumption in these models is that consumers face liquidity risk; that is, they are unsure whether they wish to consume in the short run (intermediate period) or long run (final period). Banks provide a mechanism for insuring against this liquidity risk. A bank can pool the resources of consumers and offer a demand deposit contract that yields a higher expected utility than a consumer can obtain without a bank, for both types of consumer. The banks in the literature reviewed here are all assumed to be subject to (local) competition, so that they cannot extract any surplus, but make choices to maximize the ex-ante expected utility of an arbitrary (local) consumer.

In the seminal model of Diamond and Dybvig (1983), the single bank problem is deterministic. The productive investment is risk free but fully illiquid (i.e. it cannot be liquidated in the intermediate period to finance early withdrawals). The key result is that the bank is exposed to multiple equilibria due to a sequential service assumption in the intermediate period; that is, withdrawing consumers are served in the sequence in which they lodge their claims. All early consumers necessarily withdraw in the intermediate period, but late consumers are faced with a coordination problem: if a late consumer believes that all other late consumers will withdraw early, the best response would also be to withdraw early. By contrast, if a late consumer believes that all other late consumers will not withdraw early, the best response would not be to withdraw early. A bank run is therefore one of the Nash equilibria of the coordination game among consumers. In Diamond and Dybvig (1983), there is no way to determine which equilibrium will occur, but policies like deposit insurance can make the no-run equilibrium unique. Ennis and Keister (2006) add a signal to the model which consumers use to coordinate their decision to run. They are thus able to characterize the probability of a run uniquely.

In much of the rest of the literature reviewed below, as in my model, the possibility of a run induced by a coordination problem is ruled out by assumption, in order to focus on other types of risk to the bank; for example, aggregate liquidity uncertainty and/or investment uncertainty.

I build directly on the model of Allen and Gale (2000). They study a situation with four regional banks which are subject to perfectly negatively correlated regional liquidity demand shocks, augmented with an unanticipated (zero-probability) aggregate liquidity demand shock. The productive investment is risk-free and partially liquid; that is, it can be liquidated early (in the intermediate period) but at a negative net return. Runs due to coordination failure are ruled out, so that all bank runs are efficient.

The first step in their analysis is to characterize the efficient benchmark. They show that the first best is deterministic and can be feasibly decentralized by each bank holding a sufficient proportion of its assets in the demand-deposit contracts of the other three banks. The regional liquidity demand variation serves as incentive for banks to mutually insure via cross-holdings of the demand deposit contracts. In section 5.1, I show the equivalent of this allocation for two banks.

In this decentralized allocation, contagion can occur if an unanticipated (zero ex-ante probability) aggregate liquidity demand shock occurs. Contagion occurs if (i) the aggregate liquidity demand shock is large enough, and (ii) the network of direct balance sheet links is incomplete enough. I add to these results by showing that, in a two-bank setting, contagion can still occur even when the aggregate liquidity demand shock has positive probability ex ante, and when the interbank network is complete, i.e. when each bank holds some of the demand deposit contract of the other. In my model a similar requirement for contagion arises as in Allen and Gale (2000) in the sense that the aggregate liquidity demand shock must be large enough to induce contagion.

The next closest model to the one in this dissertation is that of Castiglionesi (2007), who also builds directly on the four-bank model of Allen and Gale (2000). He studies whether a central bank setting reserve requirements can avoid contagion when an aggregate liquidity demand shock occurs with positive probability. The productive investment is risk free and partially liquid, as in Allen and Gale (2000). As in my work, Castiglionesi allows for an aggregate liquidity demand shock with positive probability, but this shock can hit only one bank. His key result is to show that the central bank can indeed avoid contagion by imposing a sufficiently high liquidity holding requirement. This result requires two features: (i) the central bank must be allowed to offer state contingent consumption when the aggregate liquidity demand shock occurs, and (ii) the regional banks must not anticipate the aggregate liquidity demand shock while the central bank must anticipate it with correct probability. Finally, when the central bank is constrained to offer deposit returns that are not state contingent, the first best allocation cannot be implemented, but contagion can still be avoided.

An important feature of the Castiglionesi (2007) model is that the aggregate liquidity demand shock can hit only one region. As such there is no ex-ante reason for banks to consider mutually insuring against the aggregate risk.

My model contributes several features and results to that of Castiglionesi (2007). First, I consider a situation where the aggregate liquidity demand shock can hit any region with the same, known probability, implying a symmetric ex-ante problem for banks in the decentralized setting, and one in which there is an incentive to potentially insure against even the aggregate liquidity demand shock. Second, I study a situation in which banks always have an accurate belief of the probability of the aggregate liquidity demand shock. Thus, my results do not rely on an inaccurate understanding of the aggregate liquidity demand risk. Third, I study non-cooperative choices over interbank deposits whereas the results in Castiglionesi (2007) rely on a competitive interbank market (as opposed to the fully strategic problem in this dissertation). Last, I show that it can be optimal to allow contagion, whereas Castiglionesi (2007) imposes that the efficient allocation must avoid contagion. In my model, with a more complete benchmark, if a central bank were to avoid contagion by imposing sufficiently high liquidity requirements on regional banks, it would in fact be welfare decreasing. This comes from my result that contagion can be optimal as in Allen and Gale (1998).

In the above models, contagion occurs in a financial system that is comprised of only banks. A question that arises is whether augmenting the model with a competitive financial market could avoid contagion. This is studied by Allen and Gale (2004*a*,*b*), who consider the interaction between a system of competitive banks and a financial market. The productive investment is risk free and can be traded on a competitive secondary market. They show that, if the financial market for aggregate risks is complete, then the full system implements the constrained efficient equilibrium. If the market for aggregate risk is incomplete, there remains a role for regulating liquidity provision. In an extension, Allen and Gale (2004*a*) show that this setup implies financial fragility – even vanishingly small aggregate liquidity demand shocks are sufficient to induce large financial crises.

My results show that, to be fully general, an aggregate banking benchmark must include the option for a global bank to consider choosing to default (the benchmark equivalent of contagion in the decentralized case). This is related to the work of Allen et al. (2009), who study a situation with a competitive interbank market where there are also only idiosyncratic and aggregate liquidity demand uncertainty as in this dissertation, but they constrain their analysis to levels of uncertainty small enough not to induce any bank failure (or contagion). The interbank market consists of the partial sale of the investment project in the intermediate period in a competitive market, where the equilibrium price of the secondary market for investment is the market clearing mechanism. They show that, in the presence of aggregate liquidity demand shocks, the interbank market fails to attain the constrained efficient allocation, as price volatility is needed to clear the market for liquidity provision in exchange for the productive investment. In their model, a central bank can implement the constrained efficient allocation by fixing interest rates.

My model nests their constrained efficient allocation as a special case. The competitive interbank market with low prices for the productive investment functions like the penalty on early liquidation of investment in my model. Since this is a purely market driven effect in their model, the aggregate benchmark in Allen et al. (2009) is not subject to this penalty. As such, the constrained efficient allocation in their paper does not employ any liquidation of the productive investment. My work therefore contributes a new benchmark that adds the feature that the early partial liquidation of productive investment, or even default with full liquidation of all assets, may be optimal, even when the benchmark is also subject to a penalty on early liquidation. In other words, I characterize an aggregate benchmark in a more general setting.

My choice of a strategic interbank market where banks invest directly in each other's demand deposit contracts, rather than a competitive interbank market that is distinct from the demand deposit contract, avoids the problem of multiple equilibria found in Freixas et al. (2011). They study a model where banks trade on a competitive interbank market with an endogenous interbank interest rate that is distinct from the deposit returns that banks offer to their consumers. The productive investment is risk free and illiquid. In their model there are two states: one in which deposit demand is identical across banks, and a second where some banks have low demand and others high demand. The key results are that there is a continuum of possible equilibria across a range of equilibrium interbank interest rates and that the central bank that selects the appropriate, state-contingent interest rate can implement the first best outcome (in the absence of aggregate liquidity demand risk) or a Pareto-improved outcome (in the presence of aggregate liquidity risk).

1.3 Models with liquidity and investment risk

The models discussed in the previous section consider only pure liquidity risk in order to isolate these effects from other causes of bank distress. The next set of models I review are also close to mine, but add investment risk as an additional cause of contagion. Thus, it is not possible to disentangle the impact of pure liquidity risk. Nevertheless, the analytic approaches and results in these models are sufficiently close to mine to deserve detailed treatment.

As in my aggregate benchmark, Allen and Gale (1998) find that contagion can sometimes be optimal. They study the outcomes for a representative bank against several benchmarks. In their model there is no aggregate liquidity risk, but there is aggregate risk in the productive investment, which they interpret as the effects of business cycles on bank value. They show that, in a setting where banks are constrained to offer a deposit return that is non-contingent in the absence of a run (or default), bank runs are efficient. In their model, they consider a representative bank, so could not study contagion. My result extends this optimal default outcome to a situation with only liquidity demand risk. The result also requires the constraint on banks that early consumption must be non-state-contingent (or risk free) in the absence of default.

My parameterization of the model includes all possibilities, from fully illiquid to fully liquid investment. I show that the constraint of a non-contingent deposit return in the absence of default is sufficient to imply the optimality of global bank default (which is the aggregate benchmark equivalent of contagion in the decentralized setting) when a large and rare enough aggregate liquidity demand shock is realized and when the early liquidation return on the investment is low enough. I show that when contagion occurs in a non-cooperative, decentralized setting, it remains as efficient as in the aggregate benchmark (within numerical accuracy).

One of my central contributions is to study the full strategic problem between two banks. In a similar vein, Dasgupta (2004) considers a model with two regional banks, but uses a different equilibrium concept and a different type of shock that leads to contagion. In his model, banks also choose cross holdings of deposits ex ante in order to insure against a regional liquidity demand shock, but the productive investment option available to each bank is region specific and risky. The equilibrium concept used is based on a coordination mechanism in global games described in Morris and Shin (2003), where an imperfect signal is received in the intermediate period on the size of the risky return on investment. If the signals are low enough, contagion occurs[4]. Combined with a carefully chosen timing assumption on the arrival of shocks and signals, the global games coordination mechanism allows Dasgupta (2004) to solve for a unique equilibrium, in terms of the size of signals, that determines the probability of contagion. My work differs in that I study how pure liquidity shocks can lead to contagion, by considering a pure-strategy Nash equilibrium without signals in the intermediate period, and I do not need to make a possibly delicate intra-period timing assumption.

One of my key results is that a global bank allocation can be less efficient than the decentralized case. This mirrors a similar result found by Wagner (2010) in a model of investment risk. In his model there are two regional banks, each with a local activity. Banks have the option of investing in only the local activity, or diversifying into both activities. He shows that full diversification is not optimal and that the equilibrium level of diversification chosen by individual banks is also greater than optimal. The reason for this is as follows: while diversifying reduces the probability of failure at a single bank, it makes the banks more similar and hence more likely to fail due to a systemic crisis. He extends the model to show that the same effect holds for the

[4]See also Goldstein and Pauzner (2005).

trading of interbank deposits to cover withdrawal demand. In his model, the interbank contract is structured such that banks agree to exchange sufficient resources to pay depositors in the intermediate period. My model differs in that the interbank deposit is chosen non-cooperatively in the initial period. I show that this is sufficient to avoid inefficient interbank links. In my model, for the range of parameters where full insurance would be inefficient, banks ex ante choose a lower interbank deposit so that they do not experience contagion or a systemic crisis. In this region my results are similar to that of Wagner (2010) in that full mutual insurance would be inefficient, but they differ in that banks strategically choose portfolios such that they avoid inefficient full mutual insurance in Nash equilibrium.

Finally, in my results, as in those of Allen and Gale (1998), when contagion occurs, it is optimal. This is not always the case: Freixas et al. (2000) consider a situation where contagion is possible as one of multiple equilibria, where technically solvent banks face a coordination problem on whether to honour credit lines. These credit lines operate similarly to the interbank deposits in my model but are in response to a spatial uncertainty - consumers face no liquidity risk (all consumers can consume in the final period of the model) but are uncertain about which region they will need to consume in.

Additionally there is investment risk because the return on investment differs across regions. The interbank market in terms of credit lines allows the set of N banks to provide liquidity insurance - banks with high local demand draw on credit lines from banks that have low local demand. The key mechanism in their model is the multiplicity of equilibria – there always exists an equilibrium where the interbank market vanishes (when bank renege on their agreements in the intermediate period), leading to contagion, even when all banks are technically solvent. They show that a central bank that functions as a lender of last resort, can implement the efficient equilibrium by offering central lines of credit. In their model, this is without cost as the mere promise of the central bank to provide lines of credit is sufficient for the multiplicity of equilibria to vanish so that just the efficient equilibrium remains, in which the interbank market functions perfectly. Thus, the central bank promises are without cost - in the efficient equilibrium they are never used.

My approach differs in that there is temporal rather than spatial uncertainty about consumption, as in Allen and Gale (2000, 1998), and only liquidity risk, rather than a combination of liquidity and investment return risk. Lastly, when contagion occurs in my model it is efficient, as in Allen and Gale (1998), so that there is no role for a lender of last resort without imbuing that policy entity with the ability to generate more real resources (i.e. a zero-cost lender of last resort will not change my equilibrium outcomes).

1.4 Other mechanisms that can induce contagion

In the papers reviewed above, contagion (where studied) occurs due to direct balance sheet links between banks. Contagion occurs when the failure of one bank reduces the value of the direct claims on that bank to such an extent that another bank – one that holds a sufficient amount of such claims in a failing bank – also fails, because of the reduction of value of the claims.

Several mechanisms other than direct balance sheet links have been studied as potential causes of contagion (or banking crises). I briefly touch on the most important of these channels of contagion in this section.

The channels of contagion distinct from direct balance sheet links can be summarized as follows:

(i) **Contagion due to investment in correlated illiquid assets.** All banks invest in potentially illiquid projects (e.g. project-specific loans, where only the bank in question has the skills to monitor the projects invested in (Diamond and Rajan, 2001)). When some banks invest in dissimilar projects from others, they may be less capable of accessing affordable interbank financing from the general corpus of banks. Thus, banks have an incentive to invest in similar projects so that mutual monitoring capabilities yields lower costs of access to interbank financing. This, however, exposes banks to similar risks and increases the likelihood of contagion. I present some results in the literature on this channel in the first part of this section.

(ii) **Contagion due to fire sales of assets.** All banks hold liquid assets against mostly illiquid loan investments to deal with unexpected liquidity demands. When these assets must be liquidated by distressed banks in order to accommodate liquidity demands, the general price of these assets can fall dramatically. Other banks, not facing large liquidity demands, may then experience a deterioration of their asset positions if they price assets on their balance sheets on a marked-to-market basis. Thus, the fire sale of commonly held liquid assets by distressed banks may reduce the balance sheet positions of initially non-distressed banks to such an extent that they are also at risk of failing. The second part of this section consider contributions on this channel of contagion.

(iii) **Contagion due to correlated information across different banks in a system.** Banks that invest in local projects that they have a comparative advantage in monitoring, are necessarily opaque to the rest of the system and the outside investor. Nevertheless, information about distress in one part of a system may convey information about the health of the entire system. Thus, the effects of information on one part of the system may have consequences for the rest of the system. Contributions to the literature along these lines are considered in the third part of this section.

1.4.1 Contagion due to correlated investments

One alternative channel for contagion across banks is the result of banks investing in similar types of illiquid assets. This mechanism is based on specialized monitoring skills. Banks that are specialized in similar types of projects, are more likely to understand the structure of counterparty banks and hence would be more willing to lend to such banks when they are in distress.

Diamond and Rajan (2005) extend their previous single bank setting in Diamond and Rajan (2001) to consider contagion that arises not only from aggregate liquidity shocks, but also from banks engaging in correlated investments. Their model consists of a system of banks engaged in entrepreneur-specific loans, where only the bank lending to that entrepreneur has the skills to extract value from the project if ended early. This effect leads to two channels of contagion that are hard to disentangle: aggregate liquidity shortages can lead to contagious solvency problems, and solvency problems across banks that lend to correlated projects can lead to aggregate liquidity shortages. The result of these entangled channels is that the appropriate response to a crisis by a central bank is not obvious. Depending on the (unidentifiable) cause of the liquidity problems, recapitalization of the most illiquid banks could cause contagion, whereas a liquidity injection may not be effective.

1.4.2 Contagion due to fire sales of liquid assets

Another channel through which contagion can occur, other than through direct balance sheet links, is through the mechanism of fire sales (Duarte and Eisenbach, 2018; Diamond and Rajan, 2011; Gromb and Vayanos, 2010; Acharya et al., 2009; Brunnermeier and Pedersen, 2008). This is distinct from the correlated investment in illiquid projects, because in the fire sale models banks hold fully liquid assets. The mechanism leading to contagion is the result of the effect on the market price of a liquid asset when many banks sell that asset simultaneously.

A fire sale refers to the large-scale sale of similar assets by distressed financial entities. This can lead to contagion of financial shocks in the following way. Suppose there is a set of banks that requires liquidity during a time when it is not available on the interbank market. This means that they will be forced to sell off assets to raise liquidity. If there is limited liquidity on the buyer side, sales of these assets (e.g. mortgage backed securities in the global financial crisis) can only take place at *fire sale* prices - i.e. at prices that are significantly below the value of the assets in 'normal times'. Thus, the market price of these assets can fall dramatically, which can affect initially non-distressed banks that practice marked-to-market pricing of assets. As the price of the assets sold by distressed banks falls, this can harm the balance sheets of banks not currently selling these assets, which in turn may drive the initially safe banks into distress. This is a distinct channel from that studied in my work (direct balance sheet links) and from other information spillover channels. Contagion occurs because banks are exposed to the same asset classes. Duarte and Eisenbach (2018) document the empirical importance of this effect during the build-up to the global financial crisis.

The literature on fire sales suggests that the practice of marked-to-market accounting standards may have contributed to the financial crisis. This is not undisputed. Laux and Leuz (2010) show empirically that during the crisis banks used leeway allowed in accounting rules to use alternative valuation methods for assets that were hard to value during the crisis. Thus, while fire sale effects were an important mechanism of contagion during the crisis, there is limited and equivocal evidence that this was *because* of marked-to-market accounting practices. Rather, contagion due to the fire sale effect runs primarily through liquid assets that were not difficult to price during the crisis.

1.4.3 Contagion due to correlated information

Several studies have shown that contagion can occur through informational channels; that is, bad news about one bank implies a bad signal about other banks in the system. I review three salient studies here.

Acharya and Yorulmazer (2008) study the ex ante behaviour of banks that are subject to information spillovers. These banks are exposed to a common risk factor, but can choose the idiosyncratic risk of the projects they invest in. The paper shows that information contagion can increase the cost of borrowing if there is bad news about the common risk factor at other banks, and that this increase is *larger* the less correlation there is between the idiosyncratic components of the investments of different banks. This leads to an ex ante herding effect: banks tend to invest in projects with similar properties (correlated idiosyncratic components). The reason for this herding is that the increasing effect that bad news has on the cost of borrowing for a specific bank is lower when it is part of a system of similar banks. This reduces the impact of information shocks on an individual bank, but the greater similarity between bank increases the risk of systemic problems.

Allen et al. (2012) study contagion due to information links across banks that swap claims on each other's projects in a bilateral network formation game. A key mechanism in their model is that the individual asset positions of banks are unobservable, but due to network links, creditors can use aggregate information to infer the solvency of banks in the network. They show that, if financing is short term, adverse information can cause creditors to stop rolling over bank financing, leading to an inefficient contagion outcome.

Ahnert and Georg (2018) provide a nuanced view, based on the ex-ante investment choices of banks, specifically regarding different sources of risk from potential information contagion. If the source of risk is exposure to common factors, adding information contagion increases systemic risk. However, if contagion is caused by counterparty risk, adding potential information contagion channels reduces systemic risk, as banks respond to the possibility of information contagion by reducing exposure to risky counter-parts ex ante.

1.5 The role of interconnectedness

The models above consider the causes of contagion. Two other key factors that determine the extent and severity of contagion are the structure and degree of interconnectedness across banks or financial institutions. There are a variety of ways to model such interconnections: a given abstract network structure (e.g. a core-periphery model), an endogenously formed network or an empirically constructed one.

The growing literature on contagion in networks is tangentially related to my work, so I my coverage of it is brief. In these works the structure of the network stands central as both a transmission and amplification mechanism of shocks. Networks of financial institutions can be construed in a variety of ways, depending on the nature of the links: they can be in the form of direct holdings of assets/liabilities in other institutions (e.g. Allen and Gale (2000)) or in the form of direct equity holdings (e.g. Elliott et al. (2014)). I briefly review some core results of different approaches and then conclude with how these features are related to my work.

Models of contagion in networks typically employ the method of Eisenberg and Noe (2001), and the extensions by Rogers and Veraart (2013) to solve for the equilibrium payment outcomes in a given network structure after the realization of a shock somewhere in the system. The central mechanism in these models is that a shock somewhere in the system can have indirect effects on every part of the system that has some path of connection with the location of the shock; for example, the direct holders of claims against the firm where a shock realizes experience a reduction in the value of those claims if the shock induces a default. If the entities directly connected to a shocked entity (call them proximate entities) also default because of the reduction of the value of their claims, those that only hold claims on the proximate entities also face reductions in their claims. Therefore, a shock to one point in the system can spread even to entities with no direct links to the shocked point in the system.

There are several papers that survey the general results in networks (e.g. Allen and Babus (2009) and Cabrales et al. (2015)), so I will touch only briefly on some of the more prominent contributions. The first set of models take the network structure as given. These are exemplified by Elliott et al. (2014) and Acemoglu et al. (2015) (see also Gai and Kapadia (2010), Gai et al. (2011) and Roukny et al. (2018)).

Elliott et al. (2014) consider a network of mutual holdings of equity. They show that the effect of networks, in particular the degree of network connectivity, is non-monotonic. Starting from a completely disconnected network, increasing the number of connections increases the likelihood of a shock in one part of the system spreading and causing a "cascade of failures", i.e. the increasing connectedness causes an increase in the likelihood and severity of contagion. However, as the degree of connectedness increases even further, the set of mutual connections eventually begins to act as an insurance device. As the network becomes more connected than a threshold level, the exposure of any specific entity to a failing entity becomes diluted, so that fewer points in the network are sufficiently exposed to a failing entity. Thus, above a threshold, increasing connectedness decreases the spread and severity of contagion.

Acemoglu et al. (2015) also consider the impact of network structure on the resilience of the system to shocks, but focus on direct balance sheet links (the cross holdings of assets or liabilities). They also show a type of non-linearity. Below a threshold in the size and number of negative shocks, more complete networks are more resilient. In contrast, when the shocks become more numerous and larger, less completely connected networks are more stable than more complete ones. Similarly, Allen et al. (2012) study contagion due to the information effects of an endogenously formed network in the form of direct balance sheet links. They show that if the network forms unconnected parts, contagion is endogenously contained.

A second set of models asks how networks are formed endogenously (Jackson and Wolinsky, 1996; Bala and Goyal, 2000; Dutta and Jackson, 2003; Bloch and Jackson, 2006). These models typically study whether a network can be formed as a Nash equilibrium of a general strategic game, whether such a network is stable, and whether it is efficient. The main result in this extensive literature that is relevant to this dissertation is that several regularity conditions are required for a Nash equilibrium network to be stable, and that these networks are typically not efficient.

The last set of network models considers empirical networks by the using detailed transaction level data between banks. Glasserman and Young (2015) study whether the degree of interconnectedness increases losses from contagion, under various assumptions about the distributions of shocks. They use the Eisenberg and Noe (2001) approach to theoretically consider the expected losses with and without additional amplification mechanisms, over and above the degree of connectedness. They then use an empirical approach (based on the part of the European banking system that participated in a 2011 stress test) to show that pure network spillover effects are insufficient to generate large losses. Other amplification mechanisms are needed to generate large losses from contagion. These are, for example, additional exogenous costs to bankruptcy resolutions, or common asset value reductions due to fire-sales or marked-to-market asset valuation practices. This empirical result is related to my work (tangentially), as I show that contagion is a very rare equilibrium phenomenon.

The models above are related to mine in the following way. In my model there are two banks, which can (endogenously) form only one of two network types by means of direct interbank deposits in the counterparty bank: a complete network, where each bank invests in the deposit contract of the other, or a completely disconnected network where both banks choose not to invest in the other bank at all. One of my contributions is showing that in the Nash equilibrium of a non-cooperative game, the (very simple) network is always complete in equilibrium: there is no parameter set where banks choose not to hold some of the other bank's deposit contract[5]. The links are symmetric in equilibrium, but the sizes of the positions vary depending on the parameter set. Additionally, I show that such links do not necessarily lead to contagion, and I can precisely quantify the prevalence of contagion as well as the parameter regions where contagion is possible in equilibrium. My last contribution to the network formation literature is that my 'endogenous network' does not require the regularity conditions typically required in the network formation literature that follows Jackson and Wolinsky (1996). Typically in this setting, some form of monotonicity and/or continuity is required. In my model, the network connections are fully solved for, and are non-monotonic and discontinuous across the parameter space.

1.6 Summary of contributions

In the literature reviewed above I show that there are many channels through which contagion can occur in a system of linked banks. I also show that there are several gaps in the literature on contagion that occurs through pure liquidity shocks. This dissertation provides several results that close some of these gaps by means of a complete study of general benchmarks and the equilibrium of a game between two non-cooperative banks that endogenously choose to hold direct balance sheet links in each other. I summarize the contributions I make relative to the reviewed literature in this closing section.

First, in my model environment (chapter 2), the stochastic setup considers only pure liquidity risk where an aggregate liquidity demand shock can occur with any probability, but that probability is common knowledge. This allows me to contribute results relative to the closely related models of Allen and Gale (2000) and Castiglionesi (2007), whose results require unforeseen aggregate liquidity risk, or those of Allen and Gale (1998) and Dasgupta (2004), who require investment risk combined with liquidity risk.

Second, I contribute a novel benchmark that admits contagion as an optimal outcome.

[5] There are no frictions in my model. Adding sufficient frictions may lead to an equilibrium where the network is empty, as in Babus (2016).

This extends the results of Allen and Gale (1998), who require investment risk to obtain an optimal contagion outcome, as well as those of Allen et al. (2009) and Castiglionesi et al. (2017), whose efficient benchmarks do not consider contagion as a possibly optimal outcome.

Third, I solve the pure non-cooperative equilibrium in the interbank market between two fully strategic banks. This adds to the literature, which typically studies either competitive equilibria (e.g. Allen and Gale (1998, 2004*b*,*a*), Allen et al. (2009), Freixas et al. (2011) or Wagner (2010)) or uses a different equilibrium definition that requires coordination on signals about investment risk (Dasgupta, 2004). My results may therefore be more in line with such market features as bilateral over-the-counter contracts for liquidity. This also contributes tangentially to the network formation literature (Jackson and Wolinsky, 1996; Bala and Goyal, 2000; Dutta and Jackson, 2003; Bloch and Jackson, 2006), as I solve for the simplest possible network without the imposition of any ex ante regularity conditions not derived from the fundamental economic structure of the model.

Fourth, I show that a numerical approach is useful in quantifying the global ex-ante prevalence of contagion, albeit in a highly stylised abstract model. Previous studies have typically stopped at showing that the probability of contagion could be determined, without providing any indication of how likely contagion would be in their particular model (Dasgupta, 2004), or have found that contagion was just one of many possible equilibria (Freixas et al., 2000).

Finally, I present a model where there is no obvious role for policy, when comparing the outcome of the non-cooperative, decentralized equilibrium with a global bank that is constrained to offer non-contingent demand deposit returns in the absence of contagion. Indeed, I find that the decentralized allocation is superior to the aggregate benchmark for some parameter regions. This is similar to results of inefficient diversification of Wagner (2010).

Chapter 2

Environment

There are three dates $t = 0, 1, 2$ and a single divisible good for consumption and investment. Two safe, constant returns to scale technologies are universally available. Liquidity y yields a unit gross return at each subsequent date. Productive investment x at $t = 0$ yields a gross return of R at $t = 2$ and a liquidation value of r at $t = 1$, where $0 < r < 1 < R$.

There are two regions, $k = A, B$, each inhabited by a unit mass of consumers. Consumers have a unit endowment at $t = 0$ and are initially identical. At $t = 1$, they privately learn their consumption preference, an idiosyncratic liquidity shock, as in Diamond and Dybvig (1983). A fraction of consumers $\nu_{ks} \in (0, 1)$ in region k and state s value consumption at $t = 1$ (early consumers), and the remainder value consumption at $t = 2$ (late consumers):

$$U(c_{1ks}, c_{2ks}) = \begin{cases} u(c_{1ks}) & \text{w.p. } \nu_{ks} \\ u(c_{2ks}) & \text{w.p. } 1 - \nu_{ks} \end{cases}$$

where c_{tks} denotes consumption at date t in region k and state s. The period utility function $u(c)$ is twice continuously differentiable, strictly increasing, strictly concave, and satisfies the Inada conditions, $\lim_{c \to 0} u'(c) = \infty$ and $\lim_{c \to \infty} u'(c) = 0$.

Table 2.1 shows the distribution of regional liquidity demand ν_{ks}. There are four possible states $s \in \{1, 2, 3, 4\}$, similar to Table 3 in Allen and Gale (2000). In each of states 1 and 2, there is a regional liquidity shock $\varepsilon > 0$ that is symmetric and perfectly negatively correlated across regions. This regional liquidity shock generates an insurance motive across regions, for example in the form of interbank deposits. In states 3 and 4, an aggregate liquidity demand shock hits one of the regions, resulting in possible contagion across regions if such links are formed. The size of the aggregate liquidity demand shock is $\alpha \in \left(0, \frac{1-\gamma}{2}\right)$ and it occurs with probability $p \in [0,1]$. As $p \to 0$, the Allen and Gale model (with two banks) is obtained as a special case of the model.

Table 2.1: Distribution of regional liquidity demand ν_{ks}.

State s	Probability π_s	Region A	Region B
1	$\frac{1-p}{2}$	$\nu_{A1} = \gamma - \varepsilon$	$\nu_{B1} = \gamma + \varepsilon$
2	$\frac{1-p}{2}$	$\nu_{A2} = \gamma + \varepsilon$	$\nu_{B2} = \gamma - \varepsilon$
3	$\frac{p}{2}$	$\nu_{A3} = \gamma$	$\nu_{B3} = \gamma + 2\alpha$
4	$\frac{p}{2}$	$\nu_{A4} = \gamma + 2\alpha$	$\nu_{B4} = \gamma$

At $t = 1$, the state is publicly revealed and consumers privately learn their preferences (early or late). My convention is that, under asymmetric information, a bank cannot observe the type of consumer withdrawing at $t = 1$. A late consumer who has deposited with a bank can therefore pretend to be an early consumer; therefore banks have to offer a contract that is incentive compatible, $c_{1ks} \leq c_{2ks}$, to prevent a run on the bank in state s.

As in Allen and Gale (2000), consumption allocations offered by banks at the interim date are not state contingent, $c_{1ks} \equiv c_{1k} \forall s$, unless default occurs in state s. Default allows for greater sharing of aggregate risk between late and early consumers, similar to the findings in Allen and Gale (1998). While Allen and Gale (1998) consider investment risk, I consider liquidity risk.

Finally, as in Diamond and Dybvig (1983), it is possible that multiple equilibria exist, e.g. a bank run due to self-fulfilling beliefs. The focus of this dissertation is on efficient (or essential) runs. Therefore, whenever multiple equilibria exist, I assume the no-run equilibrium is played, as in Allen and Gale (1998, 2000).

Chapter 3

Benchmarks

In this chapter, I present the benchmarks against which the properties of the decentralized results in chapter 4 are compared in chapter 6. I studied two benchmark allocations: a global bank allocation that is novel in the literature, and the autarkic benchmark of a consumer without access to a bank.

In the decentralized case, which is the main focus of the dissertation, there are two independent representative banks, one in each region. A global bank benchmark is construed as a single, representative banking entity with two branches, one in each region. This implies that the depositors at the two branches must be treated identically in all states. This is motivated by the real world resolution of a single banking company with branches in, for instance, two different cities in the same country. Such a bank cannot have one branch that fails and is liquidated, with the branch depositors receiving impaired returns; and another branch that survives with its depositors receiving unimpaired returns. Put differently, the global bank is construed as a single legal entity under a single resolution jurisdiction, with all investment done at the company level, not at the branch level. Therefore, if the global bank fails, it fails as a whole, and all depositors receive the same pay-out, regardless of the regional branch at which they deposited their endowment. Similarly, if the global bank survives, it survives as a whole and must provide identical withdrawal returns at both branches. This contrasts with the decentralized case where the bank in one region can fail while the bank in the other region survives.

Let c_{tks}^{GB} be the consumption allocation provided by a global bank to consumers that withdraw in period $t \in \{1,2\}$, in region $k \in \{A,B\}$, and in state $s \in \{1,2,3,4\}$. The interpretation of the global bank as a single entity that operates two regional branches implies that it is subject to the following constraints:

$$\begin{aligned} c_{1As}^{GB} &= c_{1Bs}^{GB} \equiv c_{1s}^{GB} \\ c_{2As}^{GB} &= c_{2Bs}^{GB} \equiv c_{2s}^{GB}, \end{aligned} \tag{3.1}$$

where the second equality in each line defines the notation that drops the region subscript to denote the identical treatment of consumers in different regions in any given state.

Moreover, given the fact that states 1 and 2 are symmetric, and so are states 3 and 4, this constraint implies that the problem of a global bank is reduced to a two state problem: in state L there is low per capita aggregate liquidity demand, $\nu_L = \gamma$; while in state H there is high per capita aggregate liquidity demand, $\nu_H = \gamma + \alpha$.

A global bank faces two further constraints that are commonly assumed in models of banking that follow Diamond and Dybvig (1983): (i) banks are entities that offer deposit returns (on early withdrawals in this model) that are not state-contingent, *unless the bank fails (defaults)*; and (ii) banks are typically subject to an information constraint: they cannot observe the type of consumer that withdraws early. That is, banks cannot observe whether consumers who attempt to withdraw in $t = 1$ do so because they must (because they are early consumers), or because they prefer to (i.e. late consumers who believe they will receive greater consumption by withdrawing early). Thus, in order to avoid a run (whereby all late consumers attempt to withdraw early), banks must offer contracts that are incentive compatible with a no run outcome: late consumers must prefer not to withdraw early.

The assumption that a bank must offer non-state-contingent consumption in $t = 1$, except if it fails and defaults, is common in the literature that the dissertation contributes to, although it is rarely given detailed discussion (see e.g. Allen and Gale (2000, 2004*b*,*a*), Dasgupta (2004), Allen et al. (2009)). It serves as a reduced form way of capturing the real world feature that bank deposit contracts are viewed as risk free as long as the bank survives. It can be formally stated as:

$$c_{1s}^{GB} \quad = \quad c_{1s'}^{GB} \quad \text{in all states } s, s' \text{ where the bank does not default} \tag{3.2}$$

The key impact of this constraint is that default is the only tool whereby consumption risk can be shared between early and late consumers, although at a high ex-post efficiency cost. One of my benchmark results is to show that, for some parameter ranges, this constraint is sufficient to render default by the global bank an optimal outcome when the aggregate liquidity demand shock realizes. Default by a global bank is optimal when the aggregate liquidity demand shock is sufficiently large and improbable. As there is only liquidity risk in the model in this dissertation, this extends the results of Allen and Gale (1998), who also find that default (and financial crisis) may be optimal, but they do so in a model which combines liquidity and investment risk.

Turning to the final constraint, as in Diamond and Dybvig (1983) and Allen and Gale (2000) (among others), banks are typically assumed to be subject to an information constraint: banks cannot observe the type of consumer (early or late) that withdraws in $t = 1$. To keep the analysis tractable, this is the *only* information constraint in the model. Thus, all consumers can observe the investment decisions of their bank in $t = 0$, as well as the realized state of nature in $t = 1$. As such, consumers can (at $t = 1$) infer the final consumption allocation across time induced by the investment choices of the bank and the realized state s. If this consumption allocation is characterized by $c_{2s}^{GB} < c_{1s}^{GB}$, all late consumers will pretend to be early consumers and attempt to withdraw in $t = 1$, and the bank will fail and default. Thus, in order to survive in state s, a bank must choose a portfolio and promised early consumption level that is *incentive compatible*. This means that the implied consumption allocation is such that late consumers find it incentive compatible wait until $t = 2$ to withdraw, rather than to run on the bank. This defines the final constraint on a global bank:

$$c_{2s}^{GB} \quad \geq \quad c_{1s}^{GB} \quad \text{otherwise a global bank defaults in state } s \tag{3.3}$$

In summary, the primary benchmark in the dissertation, the global bank, is constructed as a single banking company with two regional branches. As such, it is subject to three defining constraints:

(i) in any given state, the global bank must provide identical allocations in each region (equation 3.1),

(ii) the global bank must provide a non-state-contingent consumption allocation in $t = 1$ unless it defaults (equation 3.2), and

(iii) the global bank cannot observe consumer types and therefore must offer incentive compatible contracts in order to avoid a run by late consumers (equation 3.3, the information constraint).

The global bank allocation is therefore different from the typical aggregate benchmark in the information economics literature: the allocation chosen by a benevolent, omnipotent social planner. The standard, unconstrained planner allocation would be subject only to the defining characteristics of the economy: the preferences of consumers, investment technologies and possible states of nature. As such, an unconstrained planner would choose the first best (fully efficient) allocation. A constrained planner, i.e. one subject to the information constraint (equation (3.3)), would yield the second best (or constrained-efficient) allocation. I chose not to include these benchmarks in the dissertation as they are formally equivalent to the results presented by Castiglionesi et al. (2017) and would not constitute a contribution to the literature[1].

In the global bank benchmark, the information constraint (equation (3.3)) may or may not bind. Therefore, the global bank allocation is studied in two parts. First, the global bank allocation is studied *without* the information constraint. This is called the *global bank allocation under full information*. In this allocation, the bank is assumed to be subject only to constraints

[1]The setting and problem studied by Castiglionesi et al. (2017) are somewhat different from those in this dissertation, but the formal structure of their autarky benchmark is identical to that of the first and second best in the model considered here. They show that the first best is allocation is always incentive compatible, and hence the information constraint would never bind, implying that the first and second best allocations are identical. This allocation is characterized by state contingent early consumption and default is never optimal.

(3.1) and (3.2). In other words, it can observe the type of consumer and can prohibit late consumers from with drawing in $t = 1$. Equivalently, it can offer contracts that are *not* incentive compatible without inducing a run.

I show below that there are parameter regions where the global bank under full information optimally chooses non-incentive-compatible allocations, which means that the information constraint is economically important. I also show that there are parameter regions where the global bank under full and asymmetric information optimally chooses an allocation where it chooses to default when the aggregate liquidity demand shock realizes.

The second part of the analysis imposes the information constraint and is labelled the *global bank allocation under asymmetric information.* This part studies the impact of the information constraint on the allocation chosen by the global bank. I characterize the parameter regions where the information constraint is binding, as well as how this affects the optimal allocation.

To distinguish the two different versions of the global bank allocation, I denote the global bank under full information with superscript GB^F, and the global bank under asymmetric information with superscript GB^A.

I conclude the chapter by characterizing the autarkic allocation: the allocation chosen by a consumer without access to a bank, but who has access to the same investment opportunities as a bank. This allows an evaluation of the welfare improvement provided by deposit taking banks.

For each benchmark, I present the full analytical characterization accompanied by numerical examples that illustrate the many corner solutions that arise in this model.

3.1 Global bank under full information

A global bank under full information[2] observes the types of consumers. Because of free entry in the global banking market, it maximizes the expected utility of a consumer from an arbitrary region. The global bank can freely allocate *resources* across regions, but must treat consumers from different regions identically. Therefore, the global bank faces two (aggregate) states of nature which differ in the level of per capita liquidity demand (ν_s): in state H, which occurs with probability p, the aggregate per capita liquidity demand is high ($\nu_H = \gamma + \alpha$); in state L, which occurs with probability $1 - p$, the aggregate per capita liquidity demand is low ($\nu_L = \gamma$). At $t = 0$, the global bank chooses its portfolio and the consumption levels at each date, subject to the aggregate resource constraint at each date as well as constraints (3.1) and (3.2). Utility is strictly increasing, so all resource and budget constraints bind at the solution.

A key result is that the global bank may wish to default when aggregate liquidity demand is high. Default is the only tool available to the global bank that allows it to shift consumption risk from late to early consumers. This result is attributable to the defining constraint that consumption levels at $t = 1$ cannot be state contingent, unless default occurs. Default may thus be optimal because the constraint of non-state-contingent consumption at $t = 1$ is costly when the probability of the aggregate liquidity demand shock (p) is low.

Therefore, I solve the problem of the global bank under full information in two steps. First, I study the case in which the global bank does not choose to default in state H at $t = 1$. Call this the *no-default regime* and denote it with a subscript ND. In the no-default regime, the global bank is subject to the non-state-contingency constraint on early consumption.

Second, I study the case in which the global bank defaults in state H. Call this the *default regime* and denote it with a subscript D. In the default regime, the global bank is not subject to the non-state-contingency constraint on early consumption.

[2] I use *full information* and *observable types* interchangeably below. Similarly, for brevity, I sometimes refer to a global bank under full information as a *fully informed global bank* and a global bank under asymmetric as an *uninformed global bank*.

The two parts of the global bank problem under full information are subject to two distinct sets of economically motivated constraints, and they have to be treated separately. Let the value functions of the two parts of the problem be $V_{ND}^{GB^F}$ and $V_{D}^{GB^F}$. The complete problem of the global bank under full information (P1) is thus characterized by the combined value function:

$$V^{GB^F} = \max\left\{V_{ND}^{GB^F}, V_{D}^{GB^F}\right\}. \tag{P1}$$

The global bank under full information, therefore, has two types of choices (for a given parameter set): first, given the regime (default or no default), the global bank must choose its portfolio and the deposit return on early withdrawals to maximize ex-ante expected utility, subject to the constraints that define that regime. Second, given the maximum expected utility that can be attained in each regime, it must choose the optimal regime. In sections 3.1.1 and 3.1.2 respectively, I study the optimal choices within each regime; then I study the choice of regime in section 3.1.3.

3.1.1 No default after aggregate liquidity demand shock

Suppose the global bank chooses not to default in state H (denoted with ND). In this regime, the global bank is subject to the constraint that early consumption in states L and H must be equal:

$$c_{1L} = c_{1H} \equiv c_1$$

Given some non-state-contingent choice of early consumption, c_1, the bank has to optimally choose how to finance the chosen consumption allocation. Consider a given choice of liquidity, y, and the implied choice of investment, $x = 1 - y$. In state s, one of two situations may arise: if $y \geq \nu_s c_1$ there is excess liquidity of $y - \nu_s c_1 \geq 0$ which is transferred to the late consumer, which yields late consumption of $c_{2s} = \frac{y - \nu_s c_1 + Rx}{1-\nu_s}$. Otherwise, $y < \nu_s c_1$ and there is insufficient liquidity to fully cover early demand. In this situation, some of the investment must be liquidated early to finance $t = 1$ withdrawals. The required level of liquidation is $\frac{\nu_s c_1 - y}{r}$, which means a late consumer receives $c_{2s} = \frac{R}{1-\nu_s}\left(x - \frac{\nu_s c_1 - y}{r}\right)$

It is clearly optimal at $t = 0$ to hold sufficient liquidity to avoid certain early liquidation of the investment at $t = 1$, i.e. it is optimal to hold sufficient liquidity to fully cover the smallest possible early withdrawal: $y^* \geq \gamma c_1^*$. Since the early liquidation return in investment is $r < 1$, it is more efficient to use only liquidity to provide for early consumption in the state with low liquidity demand. Likewise, it is sub-optimal to hold more liquidity than required in the case of high aggregate liquidity demand, therefore $y^* \leq (\gamma + \alpha)c_1^*$. Since $R > 1$, it is more efficient to use investment to finance late consumption rather than excess liquidity. Taken together, optimality requires:

$$\gamma c_1^* \leq y^* \leq (\gamma + \alpha)c_1^*.$$

Therefore, the problem of the global bank under full information without default in state H (P1a) reduces to:

$$V_{ND}^{GB^F} \equiv \max_{\{c_1, c_{2L}, c_{2H}\}} (1-p)[\gamma u(c_1) + (1-\gamma)u(c_{2L})] + p[(\gamma+\alpha)u(c_1) + (1-\gamma-\alpha)u(c_{2H})] \tag{P1a}$$

$$\text{subject to:} \quad \begin{aligned} & x + y = 1 \\ & c_{2L} = \frac{y - \gamma c_1 + Rx}{1-\gamma} \\ & c_{2H} = \frac{R}{1-\gamma-\alpha}\left(x - \frac{(\gamma+\alpha)c_1 - y}{r}\right) \end{aligned}$$

Rather than solve problem P1a in terms of the portfolio choice of liquidity y and the choice interim-date consumption level c_1 (which affect welfare in both states), it is more illuminating and simpler to solve it in terms of the levels of excess liquidity (which is optimally allowed for only in state L), $e \equiv y - \gamma c_1 \geq 0$ and (early) partial liquidation of the productive investment (which is optimally allowed for only in state H), $\lambda \equiv \frac{(\gamma+\alpha)c_1 - y}{r} \in [0, x]$. Redefining choice variables in this way (which separates them by the states with and without the aggregate liquidity demand shock) yields monotonicity of the optimal choices in the probability of the aggregate liquidity demand shock p. Proposition 1 characterizes the optimal solution to the global bank problem without default under full information.

Proposition 1. *Suppose the global bank chooses not to default in state H. The optimal allocation with observable types is characterized by two unique thresholds of the probability of an aggregate liquidity demand shock, $\underline{p}_{ND}^{GB^F}$ and $\overline{p}_{ND}^{GB^F}$, with $0 < \underline{p}_{ND}^{GB^F} < \overline{p}_{ND}^{GB^F} < 1$. There are three cases:*

(i) For a sufficiently probable aggregate liquidity demand shock, $p \geq \overline{p}_{ND}^{GB^F}$, no partial liquidation occurs, $\lambda_{ND}^* = 0$. There exists a unique level of excess liquidity $e_{ND}^* > 0$ that increases in the probability of the aggregate liquidity demand shock, $\frac{de_{ND}^*}{dp} > 0$.

(ii) For a sufficiently improbable aggregate liquidity demand shock, $p \leq \underline{p}_{ND}^{GB^F}$, no excess liquidity is held, $e_{ND}^* = 0$. There exists a unique level of partial liquidation $\lambda_{ND}^* > 0$ that decreases in the probability of the aggregate liquidity demand shock, $\frac{d\lambda_{ND}^*}{dp} < 0$.

(iii) For an intermediate level of the probability of the aggregate liquidity demand shock, $\underline{p}_{ND}^{GB^F} < p < \overline{p}_{ND}^{GB^F}$, there exists a unique interior solution where both excess liquidity is held and partial liquidation occurs, $e_{ND}^* > 0$ and $\lambda_{ND}^* > 0$, with $\frac{de_{ND}^*}{dp} > 0$ and $\frac{d\lambda_{ND}^*}{dp} < 0$.

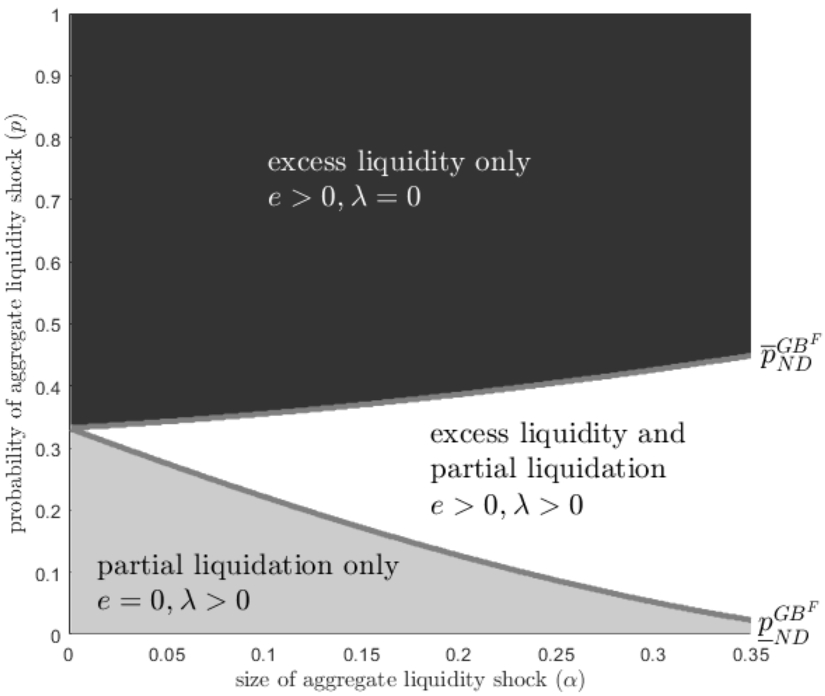
Proof. See Appendix A.1. ■

With observable types and no default in state H, the global bank uses two instruments to balance the marginal utility of consumers across states: excess liquidity and partial liquidation. Whether it is optimal to use both or only one of these instruments depends on the parameters, especially the probability (p) and size (α) of the aggregate liquidity demand shock.

Figure 3.1 illustrates the results of Proposition 1. If the global bank chooses not to default in state H, there are three regions where the solution has different characterizations: only excess liquidity in state L (for a high probability of the aggregate liquidity demand shock, $p \geq \overline{p}_{ND}^{GB^F}$); only partial liquidation in state H (for a low probability, $p \leq \underline{p}_{ND}^{GB^F}$); or both excess liquidity in state L and partial liquidation in state H (for an intermediate probability, $\underline{p}_{ND}^{GB^F} < p < \overline{p}_{ND}^{GB^F}$). The figure also shows that the size of the region in p where interior solutions occur increases in α.

To obtain intuition for the monotonicity of optimal choices in the probability of the aggregate liquidity demand shock, I consider two cases. First, when the state L is realized, any

Figure 3.1: The global bank allocation with observable types and without default in state H: the usage of partial liquidation (λ) and excess liquidity (e) varies with the size and the probability of the aggregate liquidity demand shock. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 2$, $r = 0.5$, and $\gamma = 0.3$.



excess liquidity is inefficient ex post. More resources could have been invested, while whatever level of partial liquidation chosen for state H does not affect ex post efficiency in state L. Second, when state H is realized, any partial liquidation is inefficient ex post, since more liquidity would have been desirable. Hence, an increase in the probability of state H makes allowing for excess liquidity in state L relatively more desirable than allowing for partial liquidation in state H.

The same argument does not apply to the choice of total liquidity y, which explains my focus on excess liquidity in state L and partial liquidation in state H. When p is low, no excess liquidity is held. As p increases, it becomes optimal to transfer more resources to state H. Since c_{2H} is the lowest consumption level when p is near zero, it has the highest marginal utility and is therefore the targeted consumption level to be increased. Since $e^*_{ND} = 0$ in this region, however, this happens by *reducing* total liquidity (thus increasing the productive investment), which induces a more rapid increase in c_{2H} than can be attained by only reducing partial liquidation (keeping investment constant). As p increases further, it is eventually optimal to use excess liquidity and it is possible to transfer resources to state H by *increasing* total liquidity along with reducing partial liquidation. As a result, the optimal choice of total liquidity is not monotonic. Figure 3.2 illustrates these features.

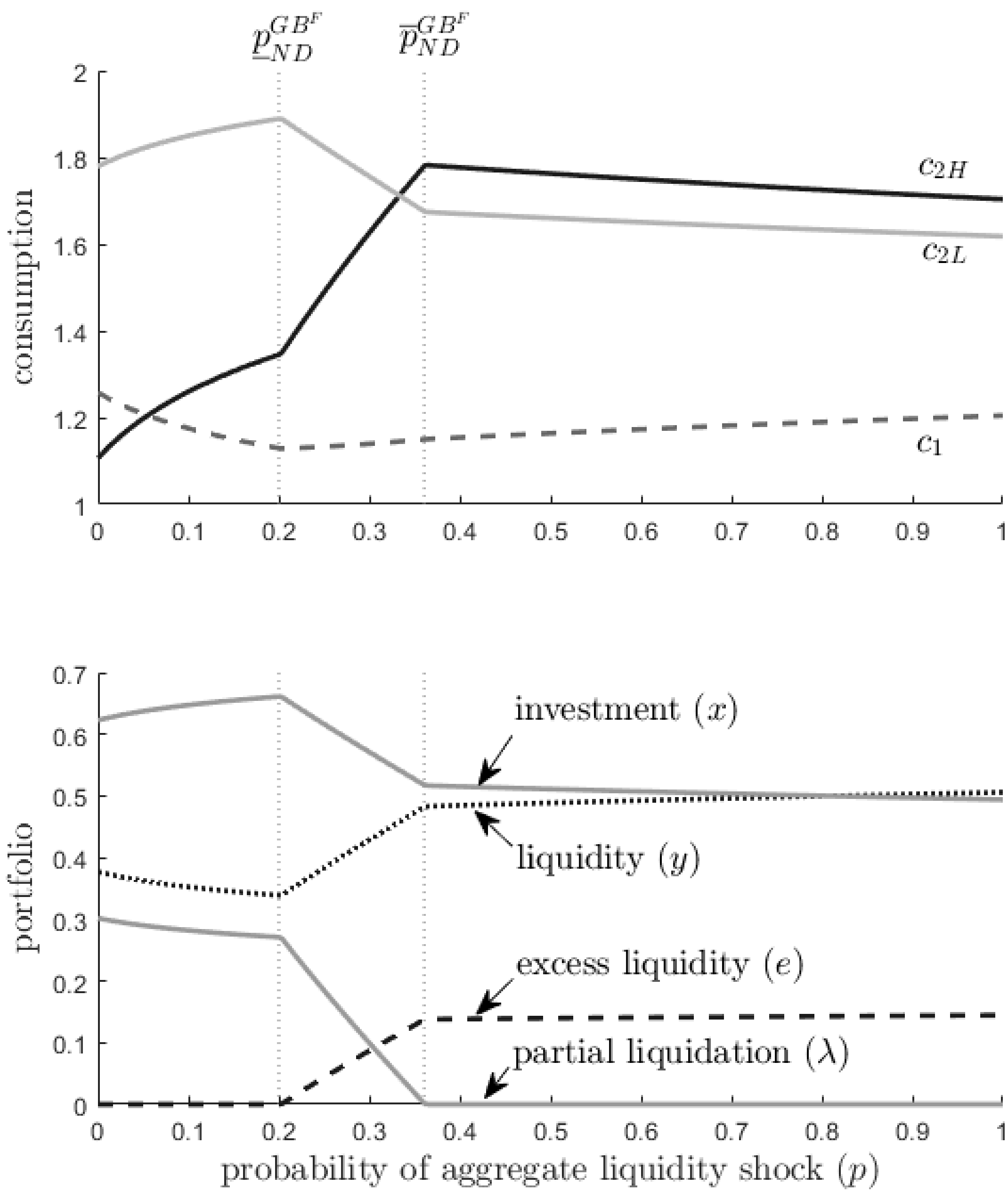
Note also in figure 3.2 that the global bank under full information chooses a consumption portfolio that is not incentive compatible in state H when p is very low.

3.1.2 Default after aggregate liquidity demand shock

Next, suppose the global bank chooses to default in state H. Since the assumed default resolution mechanism is pro-rata distribution of resources, the early and late consumers obtain the same level of consumption in state H. The implied consumption levels in state H are $c_{1H} = c_{2H} = y + rx$. Since default in state H is used only as a tool to allow the global bank to offer state contingent early consumption, and it would never be optimal to default in the state with low liquidity demand (as shown in Allen and Gale (2000)), the allocation in this regime will be characterized by $c_{2L} \geq c_{1L} \geq c_{1H} = c_{2H}$.

Again, the global bank, conditional on the choice to default in state H, must choose how to optimally finance the implied consumption allocations with its two instruments: holding excess liquidity or partially liquidating the productive investment in $t = 1$. It will never be optimal to use partial liquidation of the investment in state L, since certain liquidation can never be optimal ex ante. Since the bank defaults in state H, and hence all assets are fully liquidated, partial liquidation is no longer defined. However, it is still possible that some excess liquidity $e \equiv y - \gamma c_{1L} \geq 0$ may be held in state L, which would imply a consumption level $c_{2L} = \frac{Rx+e}{1-\gamma}$.

Figure 3.2: Global bank allocation: optimal consumption (top panel) and portfolio (bottom panel) with observable types and without default in state H: partial liquidation (λ) is non-increasing, excess liquidity (e) is non-decreasing, and the other variables are non-monotonic in p. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $\alpha = 0.12$, $R = 2$, $r = 0.5$, and $\gamma = 0.3$.



The problem of the global bank with observable types, conditional on default in state H, (P1b) reduces to:

$$V_D^{GB^F} \equiv \max_{\{c_{1L},c_{2L},c_{1H},c_{2H}\}} (1-p)\Big[\gamma u(c_{1L})+(1-\gamma)u(c_{2L})\Big]+p\Big[(\gamma+\alpha)u(c_{1H})+(1-\gamma-\alpha)u(c_{2H})\Big] \quad \text{(P1b)}$$

subject to:
$$x+y=1$$
$$c_{1L} \leq \frac{y}{\gamma}$$
$$c_{2L} = \frac{Rx+y-\gamma c_{1L}}{1-\gamma}$$
$$c_{1H} = c_{2H} = y + rx$$

Using the budget constraints in each period, I obtain an unconstrained problem in the following choice variables: the ex-ante choice of liquidity (y) and the $t=1$ excess liquidity in state L (e). The solution to this problem is characterized in Proposition 2.

Proposition 2. *Suppose the global bank chooses to default in state H. The optimal allocation with observable types is characterized by two unique thresholds of the probability of an aggregate liquidity demand shock, $\underline{p}_D^{GB^F}$ and $\overline{p}_D^{GB^F}$, with $0 < \underline{p}_D^{GB^F} < \overline{p}_D^{GB^F} < 1$. There are three cases:*

(i) For a sufficiently probable aggregate liquidity demand shock, $p \geq \overline{p}_D^{GB^F}$, all resources are kept in liquidity, $y_D^* = 1 = c_{1L}^* = c_{2L}^* = c_{1H}^* = c_{2H}^*$ and $e_D^* = 1-\gamma$.

(ii) For a sufficiently improbable aggregate liquidity demand shock, $p \leq \underline{p}_D^{GB^F}$, the global bank holds no excess liquidity in state L, $e_D^* = 0$, and there exists a unique interior solution, $0 < y_D^* < 1$ with $\frac{dy_D^*}{dp} > 0$ and associated consumption levels $c_{2L}^* > c_{1L}^* > c_{1H}^* = c_{2H}^*$.

(iii) For an intermediate probability of the aggregate liquidity demand shock, $\underline{p}_D^{GB^F} < p < \overline{p}_D^{GB^F}$, there exists a unique interior solution, $0 < y_D^* < 1$ with $\frac{dy_D^*}{dp} > 0$, and some excess liquidity held in state L, $e_D^* = y_D^*(1+(R-1)\gamma)-R\gamma$. The associated consumption levels are $c_{2L}^* = c_{1L}^* > c_{1H}^* = c_{2H}^*$.

Proof. See Appendix A.2. ∎

For a low probability of the aggregate shock (state L is the more probable), it is inefficient to allow for excess liquidity in state L, since it is costly from an ex-post perspective, and the consumption allocation is characterized by $c^*_{2L} > c^*_{1L}$. As the probability of state H increases, the expected utility places more weight on state H, so the consumption allocation in this state $c^*_{1H} = c^*_{2H} = y^*_D + r x^*_D$ must increase, which requires an increase in liquidity, y^*_D (and hence a reduction in investment x^*_D). As liquidity increases, c^*_{1L} increases and c^*_{2L} falls, eventually reaching equality. Since it cannot be efficient to allow $c_{2L} < c_{1L}$ (because of strict risk aversion) the equality $c^*_{2L} = c^*_{1L}$ arises for a sufficiently likely aggregate liquidity demand shock. As p increases further, excess liquidity is held in state L to maintain this equality. Eventually, facing default in state H becomes so probable that it is optimal not to invest at all in order to avoid costly liquidation of the investment entirely.

3.1.3 Complete problem

Finally, I describe the solution to the complete problem of the global bank with observable consumer types. The maximum utility in each of the no-default and default regimes are characterized in sections 3.1.1 and 3.1.2 above. The final choice of the global bank under full information is to choose the optimal regime, conditional on a parameter set.

Fix the parameter set. If $V^{GB^F}_{ND} \geq V^{GB^F}_{D}$, the maximum expected utility obtainable is higher under the no-default regime than under the default regime and the global bank chooses (in $t = 0$) not to default in state H. In other words, the bank chooses to be subject to the non-state-contingency constraint on early consumption. Otherwise the global bank chooses to default in state H in order to avoid the non-state-contingency constraint. The key point here is that the choice is between two distinct regimes, each subject to a different set of constraints. As such, the optimal portfolio and consumption conditional on each regime may not be continuous across regimes.

I characterize the parameter region that yields optimal default in state H in Proposition 3.

Proposition 3. *There exists a unique upper bound on the probability of an aggregate liquidity demand shock,* $\breve{p} \in [0,1)$*, whereby the global bank chooses to default in state* H *if and only if* $p \leq \breve{p}$*. Moreover, there are bounds* $(\breve{\alpha}, \breve{r}, \breve{\gamma})$ *that define a set* $\Psi \equiv \left\{\alpha > \breve{\alpha} \cap \gamma > \breve{\gamma} \cap r < \breve{r}\right\}$ *such that* $\breve{p} > 0$ *for all levels of the investment return* R *if and only if* $(\alpha, r, \gamma) \in \Psi$*.*

Proof. See Appendix A.3. ■

The intuition for the complete problem lies in the welfare cost of keeping the consumption in $t = 1$ non-state-contingent (so that early consumption is risk free), $c_{1H} = c_{1L} = c_1$. Consider the no-default regime: when the aggregate liquidity demand shock is unlikely, there is more weight on welfare in state L, and thus the average consumption levels in L are higher than in H. In this case, however, c^*_{1H} is also high (because of the non-state-contingency constraint), which can only be achieved with high partial liquidation in state H, resulting in low levels of c^*_{2H}. As a result, the allocation in state H is highly inefficient ex post. Since the utility function is concave and the allocations in the two states are linked via the non-state-contingent early consumption constraint, this ex-post inefficiency is spread across both states. Thus, the allocation in state L is also ex-post inefficient. This manifests as a lower degree of liquidity risk insurance in state L when default is not used to avoid the non-state-contingency constraint, which decouples the allocations in the two states.

The global bank has one tool for avoiding the constraint of non-state-contingency of early consumption: default in state H with equal payouts to all consumers. This tool is equivalent to forcing the global bank to accept a large degree of ex-post inefficiency in state H in order to break the link across states induced by the non-contingency of the early consumption. This allows an allocation in state L that is more efficient ex post (characterized by a greater degree of liquidity risk insurance in L than is possible without default in H). Since a low probability of an aggregate liquidity demand shock implies a high weight on L, the global bank chooses default in H for sufficiently low probabilities of the aggregate shock, $p < \breve{p}$, with an associated reduction in ex-post inefficiency in L relative to the no-default regime.

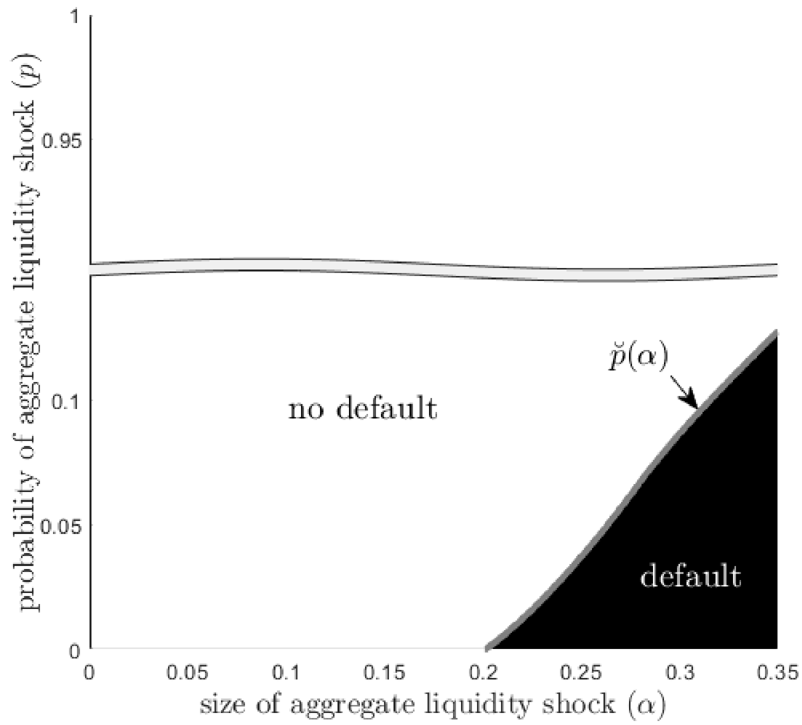
The measure of the probability of the aggregate liquidity demand shock where default is optimal depends on other parameters of the problem in an intuitive way. For a small aggregate liquidity demand shock (low α) or a low penalty on early liquidation of investment (large r), the non-state-contingency of consumption levels at the interim date is not costly to maintain[3]. There is only a small additional early demand after an aggregate liquidity demand shock, and early liquidation of investment carries only a small penalty relative to holding more liquidity. As a result, the ex-post inefficiency without default allocation is small for any given probability of the aggregate liquidity demand shock; thus there may be no positive probability for which default is optimal. Therefore, $\breve{p} = 0$. Conversely, $\breve{p} > 0$ whenever the size of the aggregate liquidity demand shock, α, is large enough and early liquidation return on investment, r, is small enough.

Figure 3.3 shows parameter regions in the probability (p) and size (α) of the aggregate liquidity demand shock where the global bank chooses to default in state H. The boundary between these regions, $\breve{p}(\alpha)$, is characterized in Proposition 3. The figure shows that the shock needs to be large and improbable for the default regime to be an optimal choice at $t = 0$. If the shock is too small, there is no positive probability for which the global bank chooses to default in state H.

Figure 3.4 shows how the consumption levels and asset portfolio choices vary with the probability of the aggregate liquidity demand shock. The key point of interest is the discontinuity in consumption levels and asset portfolio choices at the boundary between the default regime and the no-default regime.

[3]The proportion of early consumers in state L, γ, enters the problem analytically in a similar way to α, hence the impact on the solution character of γ is similar to that of α.

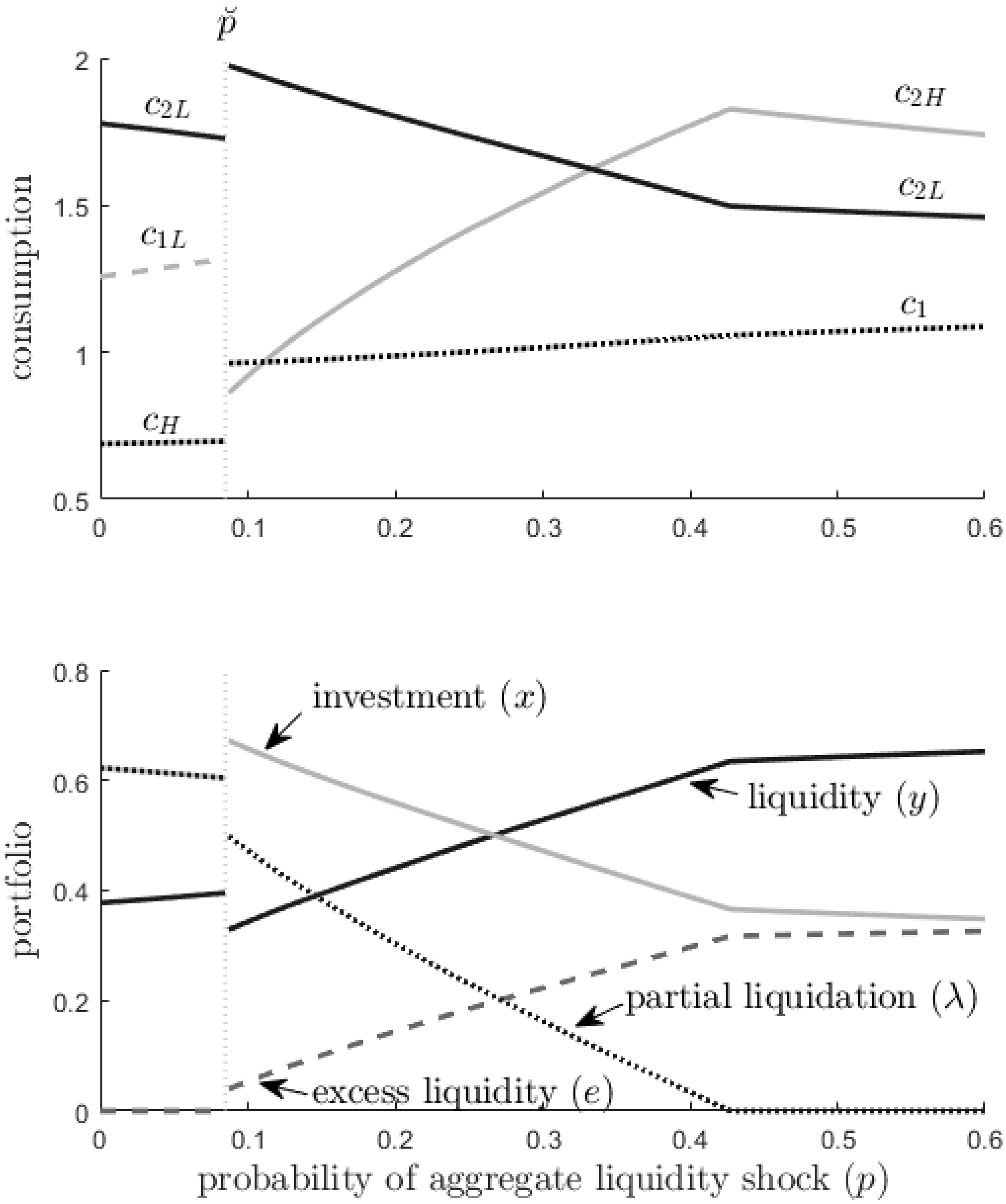
Figure 3.3: The complete problem of the global bank with observable types. Default in state H is optimal when the probability of the aggregate liquidity demand shock is low enough, and the size of the shock is large enough. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 2$, $r = 0.5$, and $\gamma = 0.3$.



3.2 Global bank under asymmetric information

A global bank that cannot observe the types of consumers at date $t = 1$ must offer incentive-compatible contracts to avoid a run by late consumers. In contrast, a global bank that can observe the types of consumers, can simply refuse to pay out to late consumers who attempt to withdraw in $t = 1$, hence is not subject to runs by assumption. I assume that the no-run equilibrium is played whenever multiple equilibria exist; therefore all bank runs are efficient as in Allen and Gale (1998, 2000).

Figure 3.4: The complete problem of the global bank: consumption (top panel) and portfolio (bottom panel) with observable types. Optimal default in state H is chosen when the probability of the aggregate liquidity demand shock is low enough, $p < \breve{p}$. There is a discontinuity in each consumption and portfolio choice at the boundary between the regions where the either the default or no-default regimes are superior. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $\alpha = 0.3$, $R = 2$, $r = 0.5$, and $\gamma = 0.3$.



In the no-default regime, this information constraint implies the following two incentive compatibility constraints on the global bank under asymmetric information:

$$c_{2H} \geq c_1, \qquad c_{2L} \geq c_1.$$

Otherwise, a global bank that cannot observe consumer types is identical to a global bank that can observe consumer types.

Because of free entry, the global bank maximizes the expected utility of an arbitrary consumer in either region. As in the full information case, the problem also has two parts: either there is default in state H, or there is not. The problem of a global bank under asymmetric information (P2) is defined by the combined, constrained value function:

$$V^{GB^A} = \max\left\{V_{ND}^{GB^A}, V_D^{GB^A}\right\} \tag{P2}$$

$$\begin{aligned} V_{ND}^{GB^A} &\equiv \left\{V_{ND}^{GB^F} \middle| c_{2L}, c_{2H} \geq c_1\right\}, \\ V_D^{GB^A} &\equiv \left\{V_D^{GB^F} \middle| c_{2L} \geq c_{1L}\right\}, \end{aligned}$$

where $V_{ND}^{GB^F}$ is defined in problem P1a and $V_D^{GB^F}$ is defined in problem P1b.

Proposition 4 characterizes the ranges of parameters where the global bank allocation under full information is incentive compatible and thus equivalent to that of a global bank under asymmetric information. That is, the inability of the global bank to observe the type of consumer is immaterial for these parameters.

Proposition 4. *The allocation of the global bank with observable types is incentive compatible in the following cases:*

(i) if $p \leq \check{p}$, that is when the fully informed global bank chooses to default in state H,

(ii) if $p \geq \hat{p}_{IC}$, where $\hat{p}_{IC} \in [0, \frac{(R-1)r}{R-r})$ is unique, and

(iii) if state L is realized.

Proof. See Appendix A.4. ■

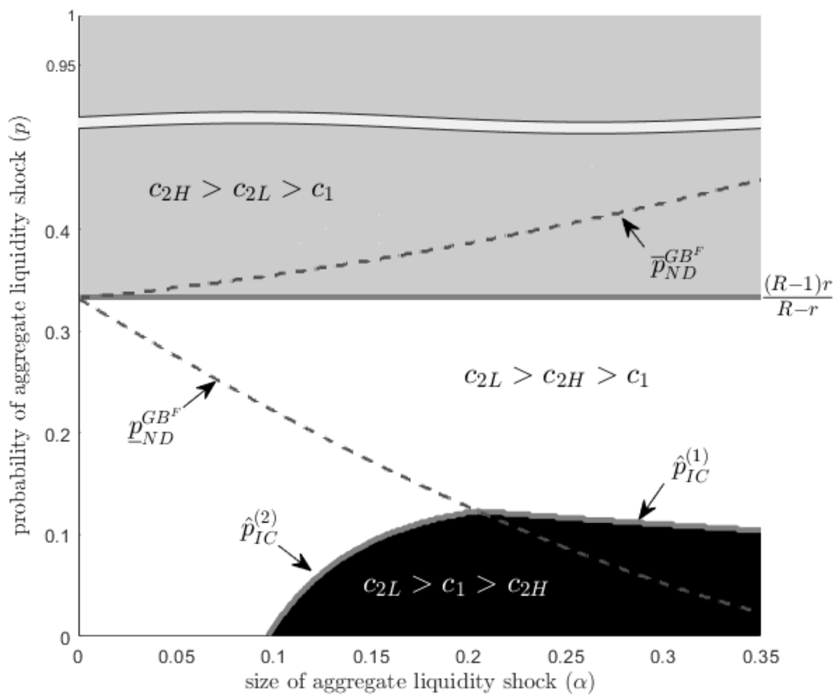
Whenever the fully informed global bank chooses to default in state H ($p \leq \check{p}$), the allocation is always incentive compatible. There are two reasons for this. First, in H, default ensures incentive compatibility by definition, $c_{1H} = c_{2H} = y + rx$. Second, given $c_{1H} = c_{2H}$, it is never optimal for $c_{2L} < c_{1L}$ as $c_{2L} = c_{1L}$ would also be feasible and is strictly preferred due to risk aversion.

Suppose the fully informed global bank chooses not to default in state H ($p > \check{p}$) (i.e. the non-state-contingency constraint on early consumption binds: $c_{1L} = c_{1H} = c_1$). If the probability of an aggregate liquidity demand shock p is low enough, the average consumption level in state L is larger than in state H; therefore $c_{2L} > c_{2H}$. Since the non-state-contingency constraint applies without default, the consumption level c_1 in state H has to be large relative to the average consumption in state H. This can only be achieved through a high amount of costly partial liquidation, resulting in a very low level of c_{2H}, which may be below c_1. Taken together, for a sufficiently low probability of the aggregate liquidity demand shock, $p < \hat{p}_{IC}$, the allocation of a fully informed global bank violates incentive compatibility, $c^*_{2H} < c^*_1$. Lastly, I obtain $\hat{p}_{IC} < \frac{(R-1)r}{R-r}$, since $c^*_{2H} > c^*_{2L}$ whenever $p > \frac{(R-1)r}{R-r}$, and the proof of Proposition 4 shows that $c_1 < c_{2L}$ always holds in state L.

Figure 3.5 illustrates the results in Proposition 4. If the fully informed global bank does not fully liquidate, there are three orderings of the consumption levels across states. The consumption allocation in state H is not incentive compatible, $c_{2H} < c_1$, if $p < \hat{p}_{IC}(\alpha)$. Otherwise, the levels in both states are incentive compatible with $c_{2L} > c_{2H} > c_1$ if $\hat{p}_{IC}(\alpha) < p < \frac{(R-1)r}{R-r}$, or $c_{2H} > c_{2L} > c_1$ if $p > \frac{(R-1)r}{R-r}$. The boundaries between these regions are characterized in the proof of Proposition 4.

Proposition 5 characterizes the optimal choice of a global bank under asymmetric information when the allocation chosen by the global bank under full information is not incentive compatible.

Figure 3.5: The global bank allocation with observable types and without default in state H: the optimal allocation in state H is non-incentive compatible if the probability of the aggregate liquidity demand shock is small enough, and the size of the aggregate liquidity demand shock is large enough. The boundary, $\hat{p}_{IC}$ is piecewise defined, above and below $\underline{p}_{ND}^{GB^F}$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 2$, $r = 0.5$, and $\gamma = 0.3$.



Proposition 5. *If $p \in (\check{p}, \hat{p})$, the consumption allocation chosen by a global bank that can observe consumer types is not incentive compatible in state H. There are two cases that characterize the closest incentive-compatible allocation of a global bank that cannot observe consumer types, separated by a unique boundary $\underline{p}_{ND}^{GB^A} \in (\check{p}, \hat{p})$:*

(i) If $\check{p} < p \leq \underline{p}_{ND}^{GB^A}$, a global bank that cannot observe consumer types chooses a portfolio with partial liquidation only $(e^* = 0)$, and the optimal allocation is:

$$\begin{aligned} y^* &= \frac{\gamma r R}{\gamma r R + \alpha R + r(1-\gamma-\alpha)} \\ c_1^* = c_{2H}^* &= \frac{y^*}{\gamma} \\ c_{2L}^* &= \frac{R(1-y^*)}{1-\gamma}. \end{aligned}$$

Moreover, $\underline{p}_{ND}^{GB^A} < \underline{p}_{ND}^{GB^F}$*, the bound below which a global bank that can observe consumer types uses only partial liquidation.*

(ii) If $\underline{p}_{ND}^{GB^A} < p < \hat{p}$*, a global bank that cannot observe consumer types chooses a portfolio with excess liquidity and partial liquidation with optimal allocation* $c_{2L}(y^*) > c_{2H}(y^*) = c_1(y^*)$ *where* y^* *uniquely solves:*

$$u'(c_{2H}(y^*)) = \frac{(1-p)(r(R(1-\alpha-2\gamma)-(1-\alpha-\gamma))+R(R(\alpha+\gamma)-\alpha))}{(1-r)R(\gamma+(1-\gamma)p)}u'(c_{2L}(y^*)).$$

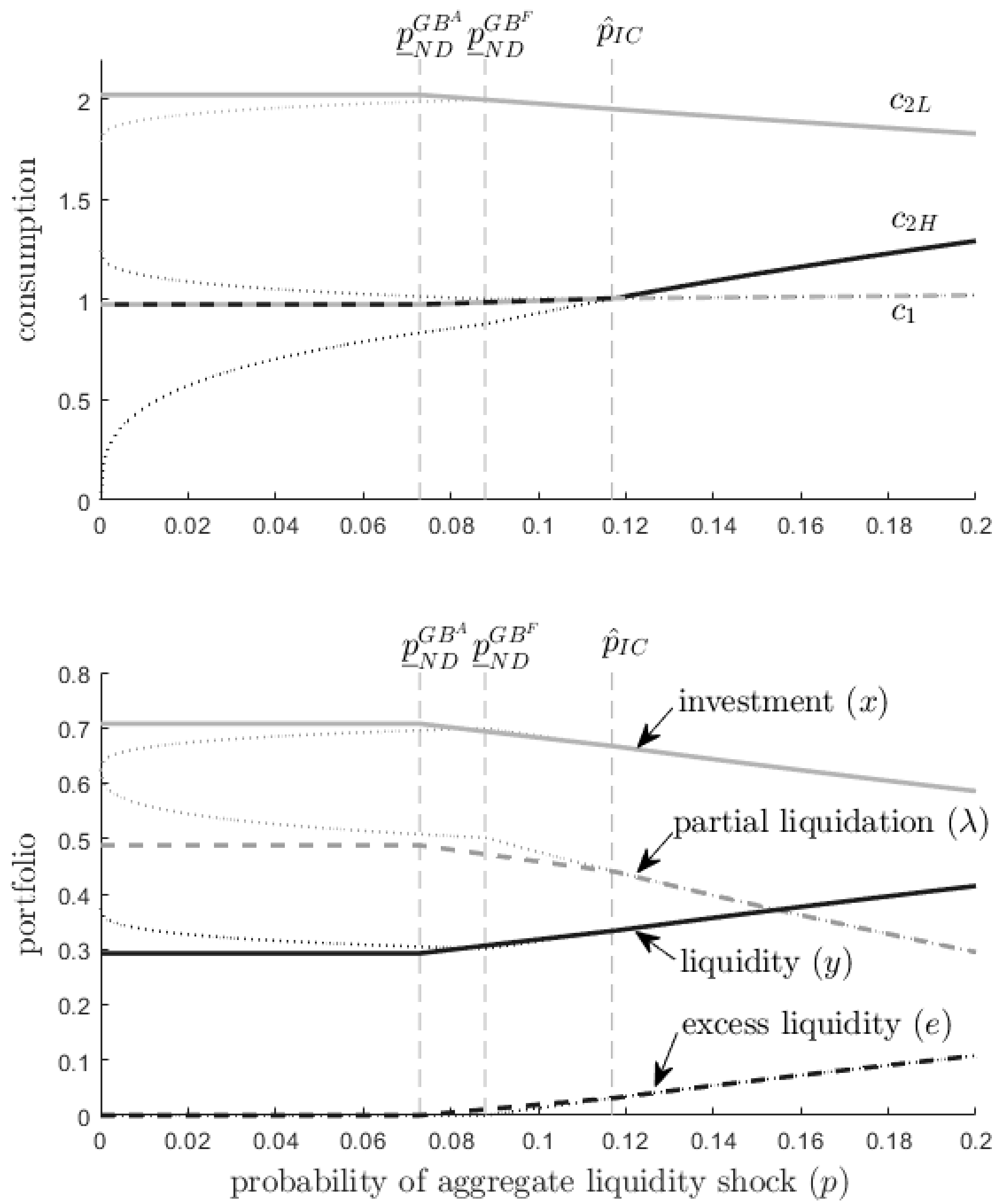
Proof. See Appendix A.5. ∎

Whenever the allocation of a fully informed global bank is not incentive compatible, an uninformed global bank optimally chooses the closest incentive-compatible allocation to that of a fully informed global bank. Since an uninformed global bank must offer $c_{2H} = c_1$ when the a fully informed global bank offers $c_{2H} < c_1$ in state H, an uninformed global bank optimally uses less partial liquidation than a fully informed global bank. For the same reasons as in the fully informed global bank case, when p is low, ex-post inefficiency in H is less costly than in L, so an uninformed global bank uses only partial liquidation in a neighborhood of $p = 0$. As p increases, it becomes efficient to start using some excess liquidity sooner for an uninformed global bank than for a fully informed global bank, since an uninformed global bank always uses less partial liquidation than a fully informed global bank:

$$\underline{p}_{ND}^{GB^A} < \underline{p}_{ND}^{GB^F}, \quad e^{GB^A} \geq e^{GB^F}, \quad \lambda^{GB^A} \leq \lambda^{GB^F}.$$

Since an uninformed global bank uses less partial liquidation in H than a fully informed global bank, it follows that $c_1^{GB^A} \leq c_1^{GB^F}$ and $c_{2H}^{GB^A} \geq c_{2H}^{GB^F}$. But $c_1^{GB^A} \leq c_1^{GB^F}$ also implies that an unformed global bank chooses lower early consumption than a fully informed global bank in L; therefore $c_{2L}^{GB^A} \geq c_{2L}^{GB^F}$. Thus, when their allocations differ, an uninformed global bank can provide less liquidity insurance to risk-averse consumers than a fully informed global bank can. Figure 3.6 illustrates these features.

Figure 3.6: Global bank allocation: consumption (top panel) and portfolio (bottom panel) with unobservable types and without default in state H. For comparison, the solid and dashed lines show the choices of an uninformed global bank, while dotted lines show the choices of a fully informed global bank. Below $\hat{p}_{IC}$, an uninformed global bank chooses $c_{2H} = c_1$. An uninformed global bank uses less partial liquidation than a fully informed global bank, and uses excess liquidity for lower probability than a fully informed global bank: $\underline{p}_{ND}^{GB^A} < \underline{p}_{ND}^{GB^F}$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $\alpha = 0.25$ $R = 2$, $r = 0.5$, and $\gamma = 0.3$.



3.3 Autarky

Without banks or financial markets, the idiosyncratic liquidity risk faced by a consumer cannot be pooled away. Consumers split their endowments between liquidity, y, and productive investment, x, at $t = 0$, before they learn their type at $t = 1$. Their consumption levels are $c_1 = y + rx$ if early and $c_2 = y + Rx$ if late. By strict monotonicity, all endowments are invested, $y^* = 1 - x^*$, and the consumption levels are a function of the productive investment only, $c_1(x) = 1-(1-r)x$ and $c_2(x) = 1 + (R-1)x$.

Since the effective ex-ante probability at $t = 0$ of being an early consumer is $p' \equiv \gamma + p\alpha$, the problem of consumers (P3) is to choose their portfolios to maximize their expected utility in autarky:

$$V^{Aut} \equiv \max_{x \in [0,1]} p'\, u(c_1(x)) + \left(1 - p'\right) u(c_2(x)) \tag{P3}$$

Proposition 6. *The optimal portfolio choice in autarky is determined by the effective probability of facing the idiosyncratic liquidity shock, p'. There are three cases:*

(i) For a sufficiently high effective probability, $p' \geq \overline{p}^{Aut} \equiv \frac{R-1}{R-r} \in (0,1)$, no investment occurs, $x^{Aut} = 0$.

(ii) For a sufficiently low effective probability, $p' \leq \underline{p}^{Aut} \equiv \frac{(R-1)u'(R)}{(R-1)u'(R)+(1-r)u'(r)} \in (0,1)$, full investment occurs, $x^{Aut} = 1$.

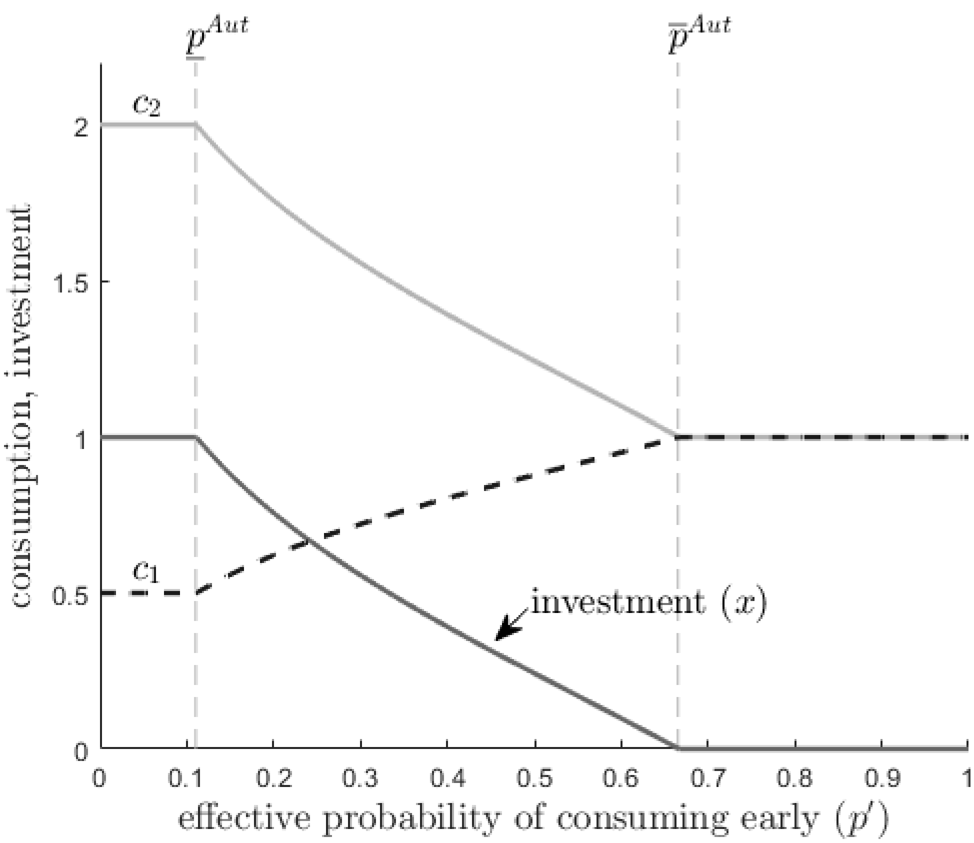
(iii) For intermediate levels, $\underline{p}^{Aut} < p' < \overline{p}^{Aut}$, there exists a unique interior portfolio choice $0 < x^{Aut} < 1$, implicitly defined by:

$$p'\,(1-r)\,u'\big(c_1(x^{Aut})\big) = (1-p')\,(R-1)\,u'\big(c_2(x^{Aut})\big)$$

Proof. See Appendix A.6. ■

When the effective ex-ante probability of being an early type is high enough, it is not optimal to invest at all ($x^{Aut} = 0$), as this would imply a high likelihood of a consumption level

Figure 3.7: Autarkic optimal consumption allocation and investment. When the effective probability of being an early consumer is low enough, $p' \equiv \gamma + p\alpha \leq \underline{p}^{Aut}$, full investment occurs and consumption allocation is maximally variable. When the probability is high enough $p' \geq \overline{p}^{Aut}$, no investment occurs and consumption is deterministic. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 2$, $r = 0.5$.



of less than unity. Storing all resources in liquidity ($y^{Aut} = 1$) guarantees $c_1^{Aut} = c_2^{Aut} = 1$. In contrast, when the effective ex-ante probability of being an early type is low enough – i.e. the risk of a consumption level of less than unity is low – it is optimal to invest fully ($x^{Aut} = 1$), as the probability of obtaining the highest possible late consumption, $c_2^{Aut} = R > 1$, is sufficiently attractive to risk obtaining the lowest possible early consumption, $c_1^{Aut} = r < 1$.

For intermediate levels of the ex-ante probability of being an early type, an interior investment is optimal, $0 < x^{Aut} < 1$. This provides insurance against obtaining the lowest possible consumption level if the consumer turns out to be an early type, at the cost of lowering the consumption level if the consumer turns out to be a late type, $r < c_1^{Aut} < 1 < c_2^{Aut} < R$. Figure 3.7 illustrates these features.

Chapter 4

Strategic Regional Banks

In each region, there is a representative bank indexed by A and B. Because of free entry within a region, each bank maximizes the expected utility of regional consumers, subject to non-negative profits. As a result, all consumers deposit their endowments at their regional bank at $t = 0$. Banks simultaneously choose the deposit return on withdrawals in period $t = 1$, d_A, a portfolio of investment and liquidity, x_A and y_A, as well as the amount of interbank deposits z_A, subject to the budget constraint at $t = 0$:

$$x_A + y_A + z_A = 1 + z_B.$$

At $t = 1$, state s realizes, consumers learn their type, all early consumers withdraw, and banks serve withdrawals. Banks choose between using liquidity, withdrawing their interbank position (denoted by w_{As}), and liquidating their investment. Late consumers withdraw at $t = 1$ if the implied contract in state s is not incentive compatible, $c_{2As} < d_A$, which forces bank default and full liquidation of all assets. Otherwise, the remaining investment matures at $t = 2$, the remaining interbank position is withdrawn, and all resources are paid out to depositors (both late consumers and the other bank). Therefore, default in state s is avoided if the implied allocation is incentive compatible, given my focus on essential runs (Allen and Gale, 1998, 2000). In the case of default, I assume pro-rata resolution with equal seniority of banks and consumers. All choices are non-cooperative.

4.1 Interbank Market and Contracts

The interbank market consists of interbank *deposits* at the other bank. As in Allen and Gale (2000), each bank treats deposits from regional consumers and the other bank identically, so the interbank deposit contract and the consumer deposit contract have the same payoffs structure.[1]

As in Allen and Gale (2000), there is a *pecking order*. A bank has a preferred sequence in which it uses its asset portfolio to serve withdrawals: first it uses liquidity, then withdraws interbank deposits, and then liquidates the investment. This pecking order is optimal for bank A as long as the relative returns between period 1 and 2 satisfy

$$1 \leq \frac{c_{2Bs}}{c_{1Bs}} \leq \frac{R}{r},$$

where c_{tBs} is the return to any depositor at bank B withdrawing in period t in state s and, therefore, $\frac{c_{2Bs}}{c_{1Bs}}$ is the inter-temporal trade-off of the interbank deposit that bank A made in bank B. The first inequality follows directly from incentive compatibility of the contract offered by bank B; the second inequality must be a feature of any deposit contract that provides ex-ante liquidity insurance. Indeed, a consumer without access to a bank also chooses a consumption allocation that satisfies this inequality, as I show in Appendix A.6.

I restrict the interbank positions to be no greater than the larger of the regional or the per capita aggregate shock: $z_{\max} = \max\{\varepsilon, \alpha\}$. This is motivated by the fact that this is the largest per capita transfer that global bank would require to support an allocation without default. Economic motivations to bound interbank positions include: (i) larger interbank holdings may be more costly than smaller ones, from an internal risk management perspective, and (ii) external

[1] There is surprisingly little empirical guidance as to the exact nature of interbank loan contracts. I assume that interbank lending is in the form of deposits, but given that about 90% of interbank loans have a maturity of only one day, the decision not to roll over an existing loan is effectively equivalent to holding demand deposits. In many countries, the vast majority of interbank loans has a maturity of only one day. For example, Arciero et al. (2016) study the euro area large value payment system Target2 and find that "From June 2008, one-day transactions (overnight, tomorrow-next, spot-next) accounted for more than 90% of total transactions". This is well in line with Furfine (2003) who studies the US fed funds market and finds that "[...] according to a Federal Reserve Bank of New York (1987) survey, overnight transactions account for 96% of the fed funds market".

capital requirements, as interbank loans, carry a positive risk weight. In practice, banks often net out large interbank positions of equal maturity.

The optimal withdrawal behaviour is as follows. A bank with the lowest regional liquidity demand keeps its interbank position, $w_{A1} = w_{B2} = 0$. A bank with the highest regional liquidity demand fully withdraws because it holds as much interbank deposits as necessary to cover this contingency, $w_{A4} = z_A$ and $w_{B3} = z_B$. In other states, each bank withdraws what is necessary to serve withdrawals from both consumers and the other bank. For bank A in state 2 (symmetric for bank B in state 1), regional consumers require $d_A(\gamma+\varepsilon)$ and demand from bank B is zero, $w_{B2} = 0$, available regional liquidity is y_A; so the shortfall is $d_A(\gamma+\varepsilon) - y_A$. The return on the deposit contract at bank B is d_B, so the total withdrawal (if positive) is $\frac{d_A(\gamma+\varepsilon)-y_A}{d_B}$. For bank A in state 3 (for bank B in state 4), the regional consumers require $d_A\gamma$ and bank B fully withdraws, $w_{B3} = z_B$; so the shortfall is $d_A(\gamma+z_B) - y_A$ and the total withdrawal is $\frac{d_A(\gamma+z_B)-y_A}{d_B}$. I summarize withdrawals in terms of choices and parameters of the model for bank A (symmetric for bank B, adjusted for the states in which regional liquidity demand is the same):

$$w_{A1} = 0, \qquad w_{A4} = z_A, \qquad w_{A2} = \frac{d_A(\gamma+\varepsilon)-y_A}{d_B}, \qquad w_{A3} = \frac{d_A(\gamma+z_B)-y_A}{d_B}. \tag{4.1}$$

The optimality of this withdrawal scheme follows from the incentive compatibility of contracts in the absence of default. The only purpose of withdrawing the interbank position at $t = 1$ is to serve early withdrawals. If local liquidity, y_A, is insufficient to cover local withdrawals at $t = 1$, the bank only wants to withdraw what is necessary to cover the shortfall. Since the fixed deposit return is d_A, a larger-than-necessary withdrawal must be held as liquidity to be paid out to late consumers in $t = 2$. Given that the contract of bank B is incentive compatible, $c_{2Bs} \geq d_B$, such withdrawals are dominated by keeping these funds invested with the other bank.

4.2 Consumption Levels

The analysis above allows me to state the levels of consumption and expected utility for any pair of bank choices. Formally, let the vector of bank choices be $\theta_k \equiv \{x_k, y_k, z_k, d_k\}$. Each pair of choices $\{\theta_A, \theta_B\}$ then implies a state-dependent consumption level for each region,

$$\{c_{1As}, c_{2As}, c_{1Bs}, c_{2Bs}\}_{s=1}^{4}.$$

Since the resource constraint at $t = 0$ binds in any equilibrium, $x_j + y_j + z_j = 1 + z_k$, I drop x_k from the vector and state it more compactly as $\theta_k \equiv \{y_k, z_k, d_k\}$. Next, I state these consumption levels in state s at bank A for two cases: without and with the default of bank A.

No default of bank A. No default requires that the implied consumption allocation is incentive compatible in all states. The total liquidity demand of bank A at $t = 1$ is $d_A(v_{As} + w_{Bs})$. This consists of the sum of the measure of early regional consumers (v_{As}) in region A, and the optimal withdrawal of w_{Bs} by bank B, at the deposit return offered by bank A, d_A. The sources of funding available to bank A to service these withdrawals are: regional liquidity, y_A; interbank withdrawals, w_{As}; and, partial liquidation of the investment, λ_{As}. Given the returns on each of these sources, the total available resources available to bank A at $t = 1$ adds to $y_A + d_B w_{As} + r\lambda_{As}$.

Depending on the total size of withdrawals relative to total available resources, bank A may end up with excess liquidity, $e_{As} = y_A + d_B w_{As} - d_A(v_{As} + w_{Bs})$, which will be paid out to late consumers and bank B at $t = 2$; or may have to partially liquidate $\lambda_{As} = \frac{d_A(v_{As}+w_{Bs})-(y_A+d_B w_{As})}{r}$ of its investment such that only $R(x_A - \lambda_{As})$ remains to be paid out to late consumers and bank B at $t = 2$. Either excess liquidity or partial liquidation may occur in some states, but never in the same state.

In sum, the consumption levels without default are:

$$
\begin{aligned}
c_{1As}(\theta_A,\theta_B) &= d_A \\
c_{2As}(\theta_A,\theta_B) &= \frac{\overbrace{e_{As}}^{\text{excess liquidity}} + c_{2Bs}(\theta_A,\theta_B)\overbrace{(z_A - w_{As})}^{\text{interbank asset}} + R\overbrace{(x_A-\lambda_{As})}^{\text{investment}}}{\underbrace{(z_B - w_{Bs})}_{\text{interbank liability}} + \underbrace{(1-\nu_{As})}_{\text{regional late consumers}}}.
\end{aligned}
$$

Default of bank A. Upon default, bank A fully liquidates all of its assets at $t = 1$. Its interbank deposits yield $c_{1Bs}(\theta_B,\theta_A)z_A$ (which is general to whether bank B also defaults or not). Its investment yields rx_A. Thus, the payout per unit of claim to all claimants, which corresponds to the *liquidation values* in Allen and Gale (2000), is:

$$
c_{1As}(\theta_A,\theta_B) = c_{2As}(\theta_A,\theta_B) = \frac{y_A + c_{1Bs}(\theta_B,\theta_A)z_A + rx_A}{1+z_B}.
$$

Therefore, for any pair of choices $\{\theta_A,\theta_B\}$, the consumption allocations in state s are:

$$
\begin{aligned}
c_{1As}(\theta_A,\theta_B) &= \begin{cases} d_A & \text{if } A \text{ does not default} \\ \frac{y_A+c_{1Bs}(\theta_B,\theta_A)z_A+rx_A}{1+z_B} & \text{if } A \text{ defaults} \end{cases} \\
c_{2As}(\theta_A,\theta_B) &= \begin{cases} \frac{e_{As}+c_{2Bs}(\theta_B,\theta_A)(z_A-w_{As})+R(x_A-\lambda_{As})}{(z_B-w_{Bs})+1-\nu_{As}} & \text{if } A \text{ does not default} \\ \frac{y_A+c_{1Bs}(\theta_B,\theta_A)z_A+rx_A}{1+z_B} & \text{if } A \text{ defaults} \end{cases}
\end{aligned}
\tag{4.2}
$$

These accounting identities are general to all possible outcomes for both bank A and bank B (no default, single default or mutual default) for any pair of strategic choices.

4.3 Nash equilibrium

A regional bank cannot observe the types of consumers at date $t = 1$, so it offers incentive-compatible contracts, $c_{2ks}(\theta_A, \theta_B) \geq c_{1ks}(\theta_A, \theta_B)$.[2] The problem of strategic bank A (P4) is to choose the regional deposit return, its portfolio, and interbank deposit (all summarized by θ_A) to maximize the expected utility of regional consumers, taking θ_B as given:

$$V_A^{SB} \equiv \max_{\theta_A} \sum_{s=1}^{4} \pi_s \big[\nu_{As} u\big(c_{1As}(\theta_A, \theta_B)\big) + (1-\nu_{As}) u\big(c_{2As}(\theta_A, \theta_B)\big)\big] \tag{P4}$$

$$\begin{aligned} \text{s.t.} \quad & x_A + y_A + z_A = 1 + z_B \\ & e_{As}, \lambda_{As} \geq 0, \\ & c_{2As}(\theta_A, \theta_B) \geq c_{1As}(\theta_A, \theta_B) \\ & \text{equations (4.1) and (4.2)}, \end{aligned}$$

where the superscript SB is used for the problem of the strategic bank, as opposed to the superscripts used for the global bank under full information (GB^F) or under asymmetric information (GB^A).

Problem P4 is symmetric across banks ex ante. It is solved simultaneously by bank k choosing θ_k in $t = 0$, taking as given the choice of the other bank. This problem defines a game $\mathcal{G}$ with a pair of symmetric best-response functions that maps a compact subspace $U \subset \mathbb{R}^3_+$ (the set of feasible choices) into itself for each bank: $\theta^{br} : U \to U$. My focus is on the symmetric equilibrium of this game.[3]

[2] This incentive compatibility constraint reads as $c_{2ks}(\theta_A, \theta_B) \geq d_k$ in all states in which the bank wishes to avoid default. A bank may choose default in some states, where the allocation is incentive compatible by construction. Thus, the incentive compatibility constraint $c_{2ks}(\theta_A, \theta_B) \geq c_{1ks}(\theta_A, \theta_B)$ is general to both cases.

[3] While asymmetric equilibria may exist in this environment, I focus on symmetric equilibria motivated by the symmetry of the problem and the symmetry of the global benchmark allocations.

Definition 1. *A symmetric Nash equilibrium of $\mathcal{G}$ is a choice, θ^*, that is a best response to itself. Equivalently, θ^* is a fixed point of the symmetric best-response function:*

$$\theta^* \quad = \quad \theta^{br}(\theta^*).$$

As in the global bank allocations, there can be more than one equilibrium type. In the decentralized case, there are three possible types of symmetric equilibria: a *no default* equilibrium (where neither bank defaults in any state), a *single default* equilibrium (where only the bank in the region where the aggregate liquidity demand shock realizes defaults), and a *mutual default* equilibrium (where both banks default whenever the aggregate liquidity demand shock realizes in either region).

4.4 Numerical Implementation

4.4.1 Motivating a numerical approach

Solving the game between two strategic banks – one in each region – requires a numerical approach because a fully analytical solution is infeasible. I expound on this claim below by considering the impact of the many potentially binding constraints in the decentralized problem relative to the few that are relevant to the aggregate problems.

Non-negativity constraints. In the no-default regime of the global bank allocation, there are two possible states. I show in section 3.1 that the non-negativity constraints on excess liquidity (optimal only in state s_L) and partial liquidation (optimal only in state s_H) induce a three-part solution (one interior and two corner solutions). Analytically, this is still manageable.

In the four-state decentralized problem, however, there are many more possible outcomes. In any state, as in the benchmarks, a bank may optimally choose to fund late consumption out of excess liquidity (if the realized liquidity demand is low relative to its expected level), or to fund early consumption by partially liquidating the investment (if the realized liquidity demand is high relative to its expected level). Consider bank A in a no-default equilibrium: it may choose to hold excess liquidity (e_{As}) in no state, in state 1 only, in states 1 and 3, or in states 1, 2 and 3. Similarly, there may be partial liquidation (λ_{As}) in no state, in state 4 only, in states 2 and 4, or in states 2, 3 and 4. It is never optimal to have excess liquidity in the state where the largest liquidity demand occurs, nor to fund the lowest liquidity demand with partial liquidation. There are thus 6 relevant non-negativity constraints:

$$\{e_{As} \geq 0\}_{s=1}^{3}, \{\lambda_{As} \geq 0\}_{s=2}^{4}$$

These non-negativity constraints may or may not bind in a number of patterns, inducing potentially 6 different characterizations of an optimal choice of bank A, given a *single* choice of bank B. The full set of possible outcomes across two banks, in the no-default equilibrium, thus has up to $6^2 = 36$ potentially different characterizations[4]. Adding the single and mutual default equilibria (with similarly numerous potential solution types, based on the results of the benchmarks) renders a fully analytic characterization infeasible.

Incentive compatibility constraints. The global bank problem under asymmetric information is equivalent to the global bank problem under full information augmented with an incentive compatibility constraint. My analytical results in section 3.2 show that this induces additional regions in the parameter set with solutions distinct from those of the global bank under full information.

Similarly, there are parameter regions for which distinct solution types exist in the decentralized problem: in any state where no default occurs, the optimal allocation may be at a point where the incentive compatibility constraint for that state, $c_{2As} \geq d_A$, does or does not bind. This means the problem has up to four potentially binding incentive compatibility constraints

[4] I do not claim that there are 36 different characterizations of the equilibrium due to these constraints, but in order to prove that a specific pattern of binding constraints is *not* an equilibrium, it would have to be given a fully analytic treatment.

for bank A:

$$\{c_{2As} \geq d_A\}_{s=1}^{4}$$

Moreover, the incentive compatibility constraints of bank A depend on the choice of bank B. This further multiplies the number of possibly distinct types of characterization required by a fully analytic approach: optimal allocations where the incentive compatibility constraint binds have a different characterization to those where it is slack. Combined with the various patterns in which the non-negativity constraints (discussed above) may or may not bind, this implies an an infeasible number of possible outcomes that a fully analytic approach must consider.

By contrast, a numerical approach can deal with all potentially binding constraints in a simple way: by imposing them on the objective function. The expected utility at any choice pair in the numerical search space that violates any constraint is numerically penalized to the extent that only choice pairs that do not violate any constraint survive the search algorithm. Thus, for any parameter set, a numerical algorithm can find the choice pairs of banks that (i) maximize ex-ante welfare, and (ii) do not violate any defining constraints, without the need of considering all possible patterns of potentially binding constraints. This makes the problem numerically tractable, where an analytic approach is not.

4.4.2 Numerically characterizing the equilibrium

I numerically solve for the symmetric Nash equilibrium of the simultaneous-move game $\mathscr{G}$, the fixed point of the symmetric best-response function. I consider a numerical approximation of this solution concept. The search space Θ of the choice variables θ is

$$\Theta \equiv \{y, z, d \,|\, y \in [0,1], z \in [0, \max\{\alpha, \varepsilon\}] \text{ and } d \in [1, R]\},$$

where the bounds on the deposit return ensure bank liquidity provision (Diamond and Dybvig, 1983) and are relevant for a utility function with relative risk aversion above unity.

Definition 2. *A numerically approximate symmetric pure-strategy Nash equilibrium of game $\mathscr{G}$ is an iteratively stable fixed point $\theta^* \in \Theta$ of the numerical best-response function $\hat{\theta}^{br}(\theta)$:*

$$\theta^* \quad = \quad \hat{\theta}^{br}(\theta^*),$$

where the iterations $\theta_i = \hat{\theta}^{br}(\theta_{i-1})$ are computed by direct numerical optimization of the constrained expected utility function over choice variables:

$$\hat{\theta}^{br}(\theta_{i-1}) \quad \equiv \quad \arg\max_{\theta_i \in \Theta} \sum_{s=1}^{4} \pi_s \left[\nu_{As} u(c_{1As}(\theta_i, \theta_{i-1})) + (1 - \nu_{As}) u(c_{2As}(\theta_i, \theta_{i-1})) \right]$$

$$\begin{aligned} s.t. \quad & x_A + y_A + z_A = 1 + z_B \\ & e_{As}(\theta_i, \theta_{i-1}), \lambda_{As}(\theta_i, \theta_{i-1}) \geq 0, \\ & c_{2As}(\theta_i, \theta_{i-1}) \geq c_{1As}(\theta_i, \theta_{i-1}), \\ & \textit{and equations (4.1) and (4.2).} \end{aligned}$$

4.4.3 Search Algorithm

The search algorithm is an iteration over the numerically constructed best-response function. In each iteration, the best response of bank A to the choice of bank B is obtained via a global search method with multiple starting points.[5]

To simplify the notation, I denote the constrained ex ante expected utility to a consumer that deposits her unit endowment at bank A from an arbitrary pair of regional choices $\theta_A, \theta_B \in \Theta$ by:

[5]Computations were performed on the Stellenbosch University's High Performance Cluster 1 (Rhasatsha): `http://www.sun.ac.za/hpc`, using the global multivariate optimization methods implemented in Matlab (2017).

$$\mathbb{E}[u(\theta_A|\theta_B)] \equiv \sum_{s=1}^{4} \pi_s \left[\nu_{As} u(c_{1As}(\theta_A,\theta_B)) + (1-\nu_{As}) u(c_{2As}(\theta_A,\theta_B)) \right]$$

$$\begin{aligned} \text{s.t.} \quad & x_A + y_A + z_A = 1 + z_B \\ & e_{As}(\theta_A,\theta_B), \lambda_{As}(\theta_A,\theta_B) \geq 0, \\ & c_{2As}(\theta_A,\theta_B) \geq c_{1As}(\theta_A,\theta_B), \\ & \text{and equations (4.1) and (4.2).} \end{aligned}$$

In numerical implementation, I set $\mathbb{E}[u(\theta_A|\theta_B)]$ to a large negative value if any constraint is violated, thus encoding the constraints directly into the objective function.

Algorithm 1. *The algorithm proceeds in the following steps (schematically represented in Figure 4.1). Fix a parameter set and let the set of equilibrium types be*

$$\mathscr{T} \equiv \{\textit{no default, single default, mutual default}\}.$$

Step 1: For each equilibrium type $\tau \in \mathscr{T}$*, perform steps 2 - 4.*

Step 2: Initialize the choice of bank B *at a feasible symmetric and incentive-compatible choice,* $\theta_0 = [y_0, z_0, d_0] \in \Theta$*. The initial point is found by means of a three-dimensional grid search of symmetric choices in the search space, with each dimension of* Θ *partitioned into* n_{grid} *points. This yields* n_{grid}^3 *potential initial choices. The potential initial choice that maximizes expected utility is chosen as the starting point for the algorithm.*

Step 3: Find the best response of bank A*,* $\theta_i = [y_i, z_i, d_i]$*, to the choice of bank* B*,* $\theta_{i-1} = [y_{i-1}, z_{i-1}, d_{i-1}]$*:*

$$\theta_i \equiv \arg\max_{\theta \in \Theta} \mathbb{E}[u(\theta|\theta_{i-1})].$$

The numerical search for θ_i *is done using standard global optimization routines with* n_{start} *randomly selected starting points. That is, for a given choice of bank B,* θ_{1-i}*, I need to find* *the best response, θ_i. I impose all inequality constraints directly on the objective function (by setting the value of the objective function to a large negative number at any pair of choices where a constraint is violated). This means that the numerical objective function has areas in its domain that are flat. Standard numerical optimization routines typically fail if they start in a flat area. To overcome this problem, I use a global search method that uses n_{start} stochastically chosen starting points (in every iteration). The algorithm then selects the final optimal θ_i as the best among the convergence points of each of the multiple starting points. Typically, all starting points that lie in an area where the objective function is not flat converge to the same point.*

Step 4: *Compute the ℓ_1 distance norm*[6] *of the difference between θ_i and θ_{i-1}: $\delta_i = \|\theta_i - \theta_{i-1}\|_1$. Replace the choice of bank B, θ_{i-1}, with the best response found, θ_i. Repeat steps 3 and 4 until one of the exit criteria is met:*

1. *I label an iteration* strongly convergent *to a symmetric equilibrium if the change from one choice to the next in the iteration falls below the convergence tolerance parameter, $\delta_i < \delta_{crit}$. This case suggests an iteratively stable numerical fixed point of the best-response function which corresponds to an equilibrium of type τ: $\{\theta_A^*, \theta_B^* | \tau\} = \{\theta_i, \theta_{i-1}\}$, with associated value function $V_A^{SB}(\theta_A^*, \theta_B^* | \tau)$.*
2. *I label an iteration* weakly convergent *to symmetric equilibrium if the number of iterations exceeds the maximum number of allowed iterations, $i > i_{max}$, with $\delta_i > \delta_{crit}$ for all i. In this case, I select the closest sequential pair as the candidate equilibrium of type τ: $\{\theta_A^*, \theta_B^* | \tau\} = \underset{\{\theta_i, \theta_{i-1}\}}{\arg\min}\, \delta_i$, with associated value function $V_A^{SB}(\theta_A^*, \theta_B^* | \tau)$*

Step 5: *Selection of equilibrium type: steps 1-4 yield a candidate equilibrium with associated value function for each of the three possible equilibrium types $\tau \in \mathcal{T}$. For a given parameter set, I select the equilibrium type τ^* as:*

$$\tau^* = \underset{\tau \in \mathcal{T}}{\arg\max}\, V_A^{SB}(\theta_A^*, \theta_B^* | \tau)$$

[6] The ℓ_1 norm is the strictest distance metric I could choose, as it corresponds to $\delta_i = |y_i - y_{i-1}| + |z_i - z_{i-1}| + |d_i - d_{i-1}|$. I choose this norm to ensure that, even in the weaker convergence criterion, the candidate equilibrium choice is as close to a symmetric equilibrium as possible.

Figure 4.1: Schematic of search algorithm.

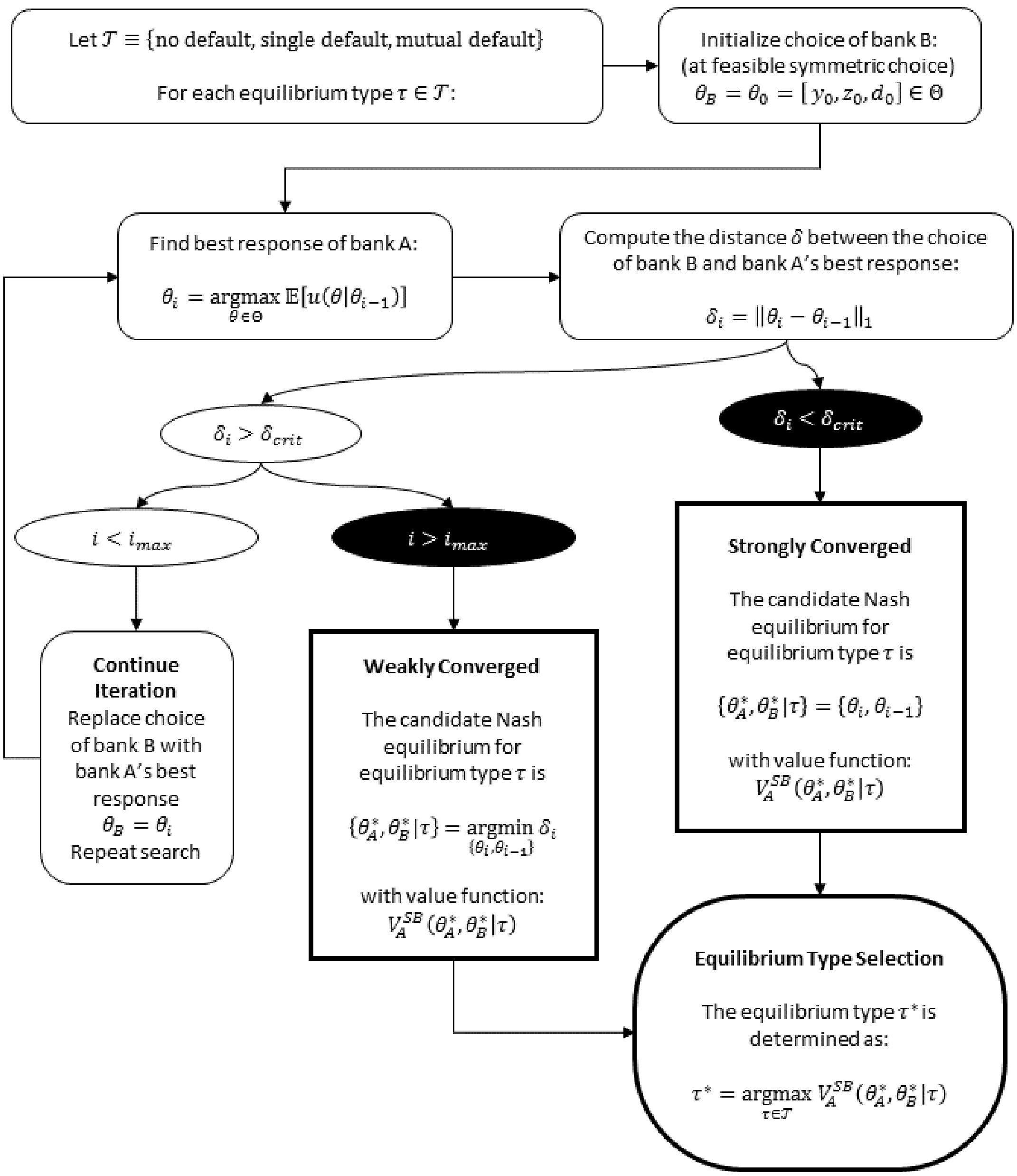


Table 4.1 summarizes the parameters that control the optimization routines.

Table 4.1: Optimization control parameters.

Parameter Name	Value
Partition of search space for initial point search (n_{grid})	25
Number of starting points for best response search (n_{start})	2500
Convergence tolerance (δ_{crit})	10^{-5}
Maximum iterations (i_{max})	$128 = 2^7$

For each parameter set, I search a grid of size n_{grid}^3 symmetric choices to find the best choice at which to *initialize* the iteration over the best-response function. Then, for each step *within* the iteration (i.e., at a given choice for bank B (θ_{i-1})), I use a global search algorithm with n_{start} stochastically selected starting points for the choice of bank A (θ_i) to find the best response. The final best response of bank A (θ_i) is the best among the convergence points of each of the stochastically selected starting points for the algorithm, given the choice of bank B.

Chapter 5

Numerical Results

In this chapter I present my numerical results for the case of regional banks, organized in five subsections, from specific to general.

The theoretical literature on financial intermediation is dominated by purely analytical work. One of the contributions of my research is to show that important theoretical insights can be gained via numerical methods applied to models that are too complicated to be amenable to comprehensive analytical characterization. An essential viability check on a numerical approach would entail demonstrating that it can replicate extant analytical results. Section 5.1 shows that the numerical approach to solving my model is able to replicate the analytic results in Allen and Gale (2000), which is a special case of my model with zero probability of an aggregate liquidity demand shock ($p = 0$).

The novel feature of the model in this dissertation is the results obtained from studying contagion over the full range of the probability (p) of the aggregate liquidity demand shock. In section 5.2 I discuss the results for contagion and the optimal portfolio choices as the shock probability varies from 0 to 1.

The analytic benchmark results in chapter 3 are focused on the interaction between the probability (p) and size (α) of the aggregate liquidity demand shock; therefore in section 5.3 I have replicated this focus numerically for the case of decentralized regional banks, also considering how the degree of risk aversion affects the prevalence of contagion.

The most important result of this research is the characterization of the prevalence of contagion in the most general way possible in a numerical approach. Thus, in section 5.4 I consider the prevalence of contagion based on a large set of random draws from the full parameter space.

Finally, to complement the results over p and α, I use machine learning techniques in section 5.5.2 to provide a comprehensive picture of how the various parameters influence the type of equilibrium, especially focusing on the values of the model parameters at which contagion can occur.

For the numerical results I focus on a standard constant relative risk aversion utility function that obeys the Inada conditions; that is, the constant elasticity of substitution utility function $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$.

5.1 Replicating the Allen and Gale (2000) results

A minimum requirement for accepting theoretical results obtained by purely computational means is that the approach is able to reproduce results established analytically. In this section I show that my solution algorithm for the decentralized banking problem replicates the results obtained by Allen and Gale (2000). Their model is a special case of mine when the probability of the aggregate liquidity demand shock is zero ($p = 0$).

Allen and Gale (2000) argue that, in the absence of aggregate risk (i.e. when $p = 0$), the first best is obtained by decentralized banks. In the absence of aggregate risk, the stochastic structure is described in Table 5.1.

Since the regional shocks cancel out when aggregated across regions, the global risk sharing problem is deterministic. Allen and Gale (2000) show that the first best arrangement satisfies the first order condition, $u'(c_1^{AG}) = Ru'(c_2^{AG})$. This allocation is feasibly decentralized as follows: each bank holds interbank position $z_A = z_B = z^{AG} = \varepsilon$ and just enough liquidity to satisfy the

Table 5.1: Distribution of regional liquidity demand ν_{ks} without aggregate shock

State	Probability	Region A	Region B
1	$\frac{1}{2}$	$\nu_{A1} = \gamma - \varepsilon$	$\nu_{B1} = \gamma + \varepsilon$
2	$\frac{1}{2}$	$\nu_{A2} = \gamma + \varepsilon$	$\nu_{B2} = \gamma - \varepsilon$

average global early liquidity demand $y_A = y_B = y^{AG} = \gamma c_1^{AG}$. If state 1 is realized, bank B needs to withdraw its full interbank holding as it faces regional consumer demand $(\gamma + \varepsilon)c_1^{AG}$ but has liquidity only sufficient to cover γc_1^{AG}. Bank A does not need to withdraw, as the sum of regional demand $(\gamma - \varepsilon)c_1^{AG}$ and demand from bank B εc_1^{AG} is equal to liquidity available locally.

If the utility function is of the standard constant relative risk aversion form $u(c) = \frac{c^{\rho-1}-1}{\rho-1}$, the optimal choices are:

$$y^{AG} = \frac{\gamma R}{\gamma R + (1-\gamma)R^{\frac{1}{\rho}}}, \quad z^{AG} = \varepsilon, \quad d^{AG} = \frac{R}{\gamma R + (1-\gamma)R^{\frac{1}{\rho}}}. \tag{5.1}$$

My approach to replicating this result proceeds as follows: set $p = 0$ and for each risk aversion parameter in the set $\rho_j = \{2, 3, 4, 5, 6, 7, 8\}$, take the following steps.

Step 1: For ρ_j, take a random draw i of the model parameters $\{R, \gamma, \varepsilon, \alpha\}$, from independent, uniform distributions.

Step 2: For each random draw i of parameters $\{R_i, \gamma_i, \varepsilon_i, \alpha_i\}$

2.1: Calculate the Allen and Gale (2000) allocation and implied expected utility, $\{EU_i^{AG}, y_i^{AG}, z_i^{AG}, d_i^{AG}\}$, using equation (5.1),

2.2: Compute the numerical equilibrium choices and equilibrium expected utility, $\{EU_i^*, y_i^*, z_i^*, d_i^*\}$, via the search algorithm presented in section 4.4.3, and

2.3: Construct the normalized results: $\{\frac{EU_i^*}{EU_i^{AG}}, \frac{y_i^*}{y_i^{AG}}, \frac{z_i^*}{z_i^{AG}}, \frac{d_i^*}{d_i^{AG}}\}$.

Table 5.2 shows the descriptive statistics of the replication exercise. While there is naturally some numerical variation, the algorithm succeeds in replicating the results of Allen and

Gale (2000). The normalized expected utility results show this most clearly: the median of the numerical results is equal to 1, as required, with the mean being only marginally below 1, and with a very small standard deviation. The skewness of the distribution of the numerical results is negative, showing that the few instances where the goal was missed can be attributed to the algorithm ending before reaching the true Nash equilibrium. The normalized choice variables show somewhat more variability, but are generally centred on unity, except for the normalized interbank position.

This last result does not imply a failure in the replication exercise, as there is a natural indeterminacy in the interbank position in symmetric equilibrium. The interbank position in the Allen and Gale (2000) allocation is equal to the size of the regional shock (ε). When the size of the aggregate liquidity demand shock (α) is larger than ε, the search space of the algorithm admits larger than necessary interbank positions. However, as long as the allocation is symmetric, a larger than necessary interbank deposit still implements the Allen and Gale (2000) allocation. If some symmetric pair of interbank positions, z^*, feasibly implements a symmetric allocation, any larger symmetric pair of interbank positions implements the same allocation. Any excess interbank holdings above what is needed to finance early withdrawals (the only purpose of the interbank position) will pay out in the second period. Given that the allocations are symmetric in terms of late payouts in the absence of aggregate risk, these second-period payouts will cancel out and have no welfare implication for consumers.

Additional evidence for the successful replication exercise is presented via scatter plots in Figure 5.1, where I plot the results of the computed equilibrium against the theoretical values in the Allen and Gale (2000) allocation. The figure shows the near-perfect correspondence of the expected utility in numerical equilibrium with the predicted analytical value (top-left panel). As with the descriptive statistics in Table 5.2, there is much greater variability in the choice variables than in the expected utility. This illustrates a central feature that complicates the numerical approach: in the region of the equilibrium, the expected utility surface is extremely flat, so that even notable variation in choices has little impact on the implied welfare of the equilibrium. Hence any numerical algorithm will struggle to converge.

The greatest variability in liquidity (the top-right panel of Figure 5.1) occurs when the analytical allocation has near-zero liquidity. This occurs when the liquidation return on investment is near 1, as then investment and liquidity are near substitutes for funding early investment and the numerical algorithm struggles to pin down the optimal level of liquidity precisely. The interbank deposit (bottom-left panel) is variable in the sense that the numerical equilibrium frequently selects (symmetric) values larger than those predicted by the analytic results. As argued above, this does not affect welfare, as interbank positions that are too large cancel out in symmetric equilibrium. Lastly, the deposit return is most variable when it is predicted to be very large, which occurs when both the early liquidation return and the return at maturity are large. Again, even this notable variability has a limited impact on implied welfare.

The final evidence of the successful replication of the benchmark is presented in the form of regression results in Table 5.3. In this analysis, the computed equilibrium values of each measure of interest were regressed on the corresponding analytical values in the Allen and Gale (2000) allocation. I used two methods: the standard ordinary least squares estimator (OLS, in the left panel of the table) and recursively weighted least squares (RWLS, right panel). RWLS is robust to outliers, which in this setting are caused by numerical errors in the computational algorithm. The OLS results show coefficients that are very close to one with very high R^2 values. In the RWLS results, all the coefficients and R^2 values are equal to one (rounded to two decimal places). In the RWLS approach I used a "fair" weighting scheme: the estimation algorithm starts with OLS (i.e. equal weights on all observations) then uses the estimated residuals to down-weight observations with large residuals (details in appendix B). This scheme down-weights outliers, but no observation received a zero weight.

I summarize the evidence in this section as a numerical result:

Numerical Result 1. *When the probability of the aggregate liquidity demand shock is zero, $p = 0$, the numerical algorithm replicates the analytic version of the results in Allen and Gale (2000).*

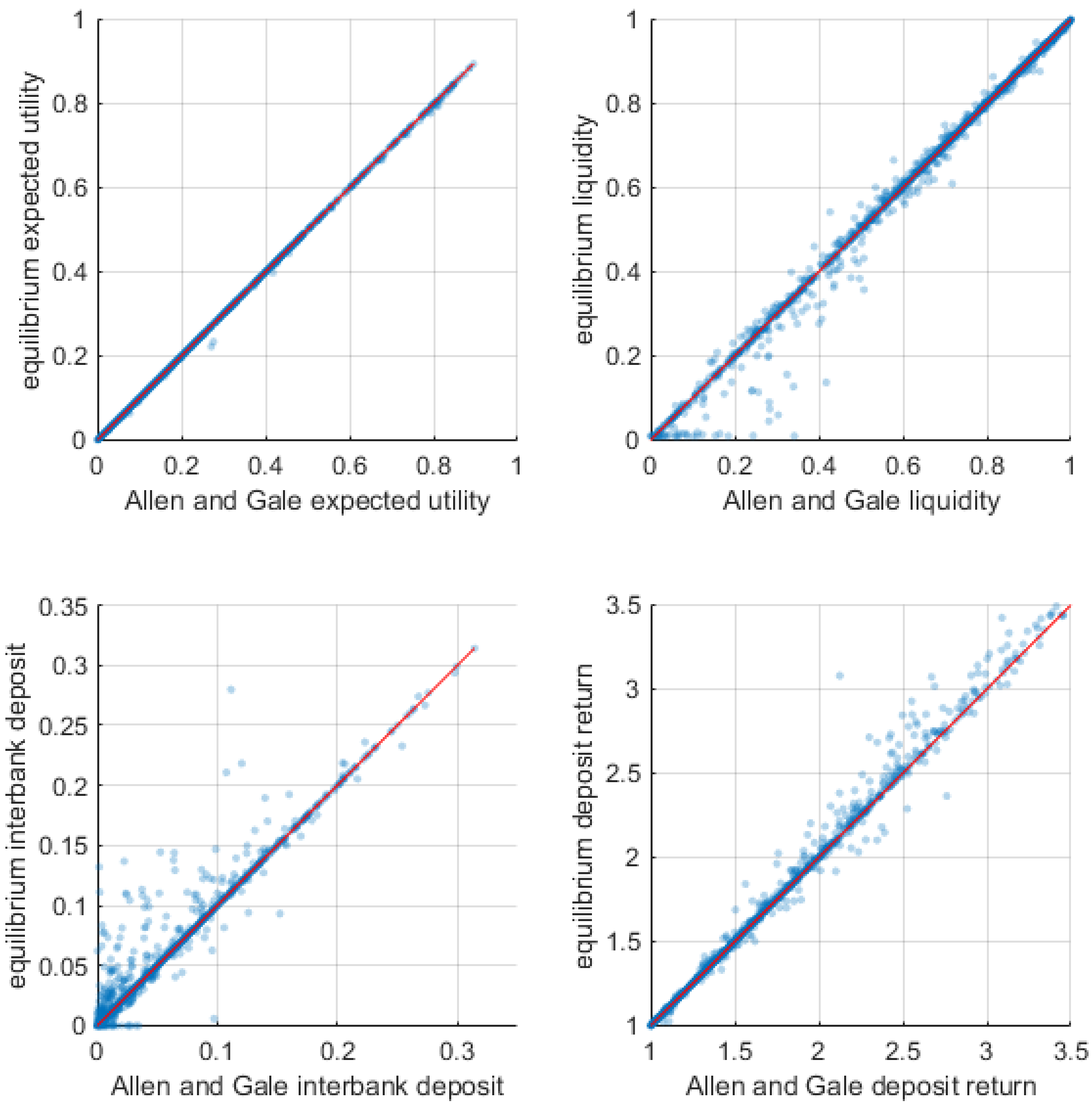
Table 5.2: Replicating the Allen and Gale (2000) result at $p = 0$: descriptive statistics of the normalized expected utility and choice variables in numerical equilibrium based on 1856 trials

Choice	Mean	Median	Std. Dev.	Skewness
Normalized expected utility ($\frac{EU_i^*}{EU_{AG,i}^*}$)	0.997	1	0.011	−10.913
Normalized liquidity ($\frac{y_i^*}{y_{AG,i}}$)	1.021	1	1.368	39.022
Normalized interbank deposit ($\frac{z_i^*}{z_{AG,i}}$)	1.662	1	8.375	26.455
Normalized deposit return ($\frac{d_i^*}{d_{AG,i}}$)	1.008	1	0.0375	7.263

Table 5.3: Replicating the Allen and Gale (2000) result at $p = 0$: regression results based on 1856 trials. For each of the values of interest, expected utility and the three elements of the bank choice variable, a pair of independent estimations were performed, regressing the computed equilibrium values on the analytical values of the corresponding variable in the Allen and Gale (2000) result. Two methods were used: ordinary least squares (OLS) and recursively weighted least squares (RWLS).

	OLS			RWLS		
	Coefficient	p-value	R^2	Coefficient	p-value	R^2
Expected utility	1.00	0.00	1.00	1.00	0.00	1.00
Liquidity	1.00	0.00	0.99	1.00	0.00	1.00
Interbank deposit	1.02	0.00	0.94	1.00	0.00	1.00
Deposit return	1.02	0.00	0.98	1.00	0.00	1.00

Figure 5.1: Replicating the Allen and Gale (2000) result at $p = 0$: scatter plots of the equilibrium expected utility and choice variable values compared with the analytical values in the Allen and Gale (2000) result based on 1856 trials. Each dot is partially transparent to show accumulation about the diagonal line from the origin, which would mean a perfect fit.



5.2 Results for the probability of the aggregate liquidity demand shock

In this section I present the first novel numerical results on the decentralized problem between two banks studied in this research. I start with results varying over only one parameter, the probability of the aggregate liquidity demand shock (p), as these results provide the strongest intuition on the nature of the different equilibrium types that emerge in the strategic game $\mathscr{G}$.

The main findings in this section are as follows. First, there are parameter values (e.g. when the aggregate liquidity demand shock is small enough) for which the no-default equilibrium is weakly superior for all positive probability levels. Thus, contagion (mutual-default equilibrium) is possible only with zero probability (where the no-default and mutual-default equilibria achieve the same expected utility). Second, there are also parameter values (e.g. when the aggregate liquidity demand shock is large enough) where the equilibrium type depends on the probability of the shock. When the probability is small enough, the mutual-default equilibrium is superior, so that contagion is possible with positive probability. For intermediate probability values, the single-default equilibrium is superior, so that only the bank in the region where the shock is realized defaults, but this does not cause contagion. Lastly, when the probability of the aggregate liquidity demand shock is large enough, the no-default equilibrium is superior, i.e. banks choose allocations such that default does not occur when the shock is realized.

My results characterize how the type of equilibrium (no default, single default or mutual default) and the equilibrium bank choices depend on the full range of probability of the aggregate liquidity demand shock $p \in [0,1]$ in two cases: when the aggregate liquidity demand shock (α) is smaller than, and when it is larger than the regional liquidity demand shock (ε). Table 5.4 summarizes the parameter sets studied in this section.

Figure 5.2 shows the expected utility for each equilibrium type when the local realization of the aggregate liquidity demand shock (2α) is small, and in particular, when it is smaller than the regional liquidity demand shock, $2\alpha = 0.02 < \varepsilon = 0.1$. The no-default equilibrium is strictly

Table 5.4: Model parameters for results across the probability of the aggregate liquidity demand shock

Parameter Name	Symbol	Value
Utility function	$u(c)$	$\frac{c^{1-\rho}-1}{1-\rho}$
Coefficient of risk aversion	ρ	2
Investment return at maturity	R	5
Investment return at early liquidation	r	0.1
Average liquidity demand	γ	0.5
Regional liquidity demand shock	ε	0.1
Aggregate liquidity demand shock size	α	0.01 and 0.15
Aggregate liquidity demand shock probability	p	$[0,1]$

superior for all $p > 0$, while the no-default and mutual-default equilibria are equivalent at $p = 0$ and attain the expected utility of the Allen and Gale (2000) allocation as in section 5.1. Thus, parameter sets exist where there is no positive probability at which contagion occurs. The intuition for this is simple. When the aggregate liquidity demand shock is smaller than the regional shock, any interbank position that is large enough to support the no-default equilibrium in the absence of the aggregate liquidity demand shock is also large enough to support the no-default equilibrium when the aggregate liquidity demand shock is realized.

In contrast, Figure 5.3 shows the expected utility for each equilibrium type when the local realization of aggregate liquidity demand shock (2α) is large and, in particular, when it is larger than the regional liquidity demand shock, $2\alpha = 0.3 > \varepsilon = 0.1$. In this case there are three regions with different equilibrium types across the probability of the aggregate liquidity demand shock. When the probability is low enough ($p \in [0, 0.06)$), the mutual-default equilibrium is superior. For intermediate probability values ($p \in [0.06, 0.18]$)), the single-default equilibrium is superior. When the probability is high enough ($p \in (0.18, 1]$), the no-default equilibrium is superior. Again, when the probability is zero, the equilibrium attains the same expected utility as the allocation in Allen and Gale (2000), within numerical precision.

Figure 5.4 shows the numerical accuracy of the results in the form of histograms across the probability of the aggregate liquidity demand shock for each equilibrium type where it is superior (these results are for the case where the shock is large, as in Figure 5.3). The top panel shows the ℓ_1 norm of the difference between the two consecutive choice vectors that are closest and constitute my computed equilibrium choice pair. The bottom panel presents the absolute value of the implied expected utility differences when each bank chooses one of the pair of equilibrium choices. The mutual-default equilibrium is accurately computed, while the single-default equilibrium is the least accurate. For the no-default equilibrium, the accuracy is lower for low values of the probability (near the boundary with the single equilibrium type). As noted in section 5.1, the variability in the expected utility differences of the equilibrium choice pairs are dramatically smaller than the distances between the corresponding choice pairs themselves. Again, this indicates that the expected utility hyper-surface is very flat around the equilibrium.

Figure 5.2: Expected utility at the approximate symmetric Nash equilibrium of the regional bank problem across the probability of the aggregate liquidity demand shock. In this parameter set, the aggregate liquidity demand shock is small relative to the regional liquidity demand shock, $2\alpha < \varepsilon$, and there is effectively only one equilibrium type (in contrast to the parameter set in Figure 5.3 where $2\alpha > \varepsilon$). The no-default equilibrium is weakly superior $\forall p \in (0,1]$. Thus, there is no positive probability at which mutual default (or contagion) occurs. However, when the probability is zero, the no-default and mutual-default equilibria are equivalent and attain the same expected utility as the allocation in Allen and Gale (2000), within numerical precision. The dots indicate where the algorithm result is strongly convergent. There are two lines for the expected utility in each equilibrium type, showing where the algorithm attained the worst convergence. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$, $\alpha = 0.01$ and $\varepsilon = 0.1$.

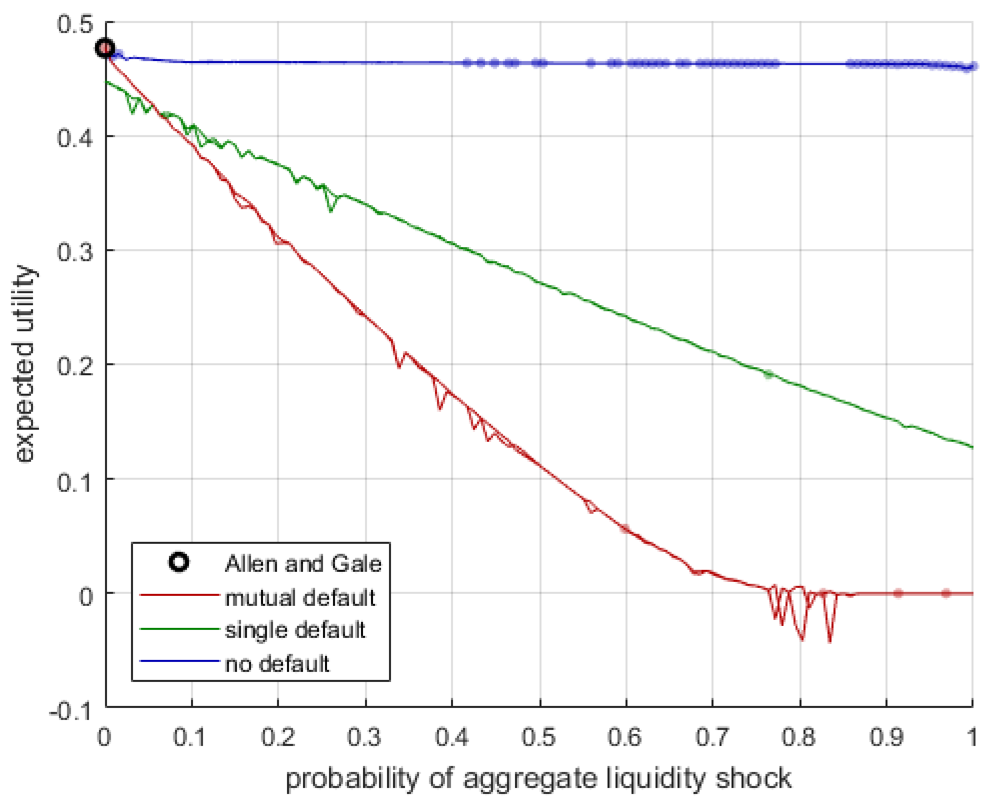


Figure 5.3: Expected utility at the approximate symmetric Nash equilibrium of the regional bank problem across the probability of the aggregate liquidity demand shock. In this parameter set, the aggregate liquidity demand shock is large relative to the regional liquidity demand shock, $2\alpha > \varepsilon$, and yields three probability regions with different equilibrium types (in contrast to the parameter set in Figure 5.2 where $2\alpha < \varepsilon$). When the probability is low enough ($p \in [0, 0.06)$), the mutual-default equilibrium is superior. For intermediate probability values ($p \in [0.06, 0.18]$), the single-default equilibrium is superior. When the probability is high enough ($p \in (0.18, 1]$), the no-default equilibrium is superior. Moreover, when the probability is zero, the equilibrium attains the same expected utility as the allocation in Allen and Gale (2000), within numerical precision. The dots indicate where the algorithm result is strongly convergent. There are two lines for the expected utility in each equilibrium type, showing where the algorithm attained the worst convergence. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$, $\alpha = 0.15$ and $\varepsilon = 0.1$.

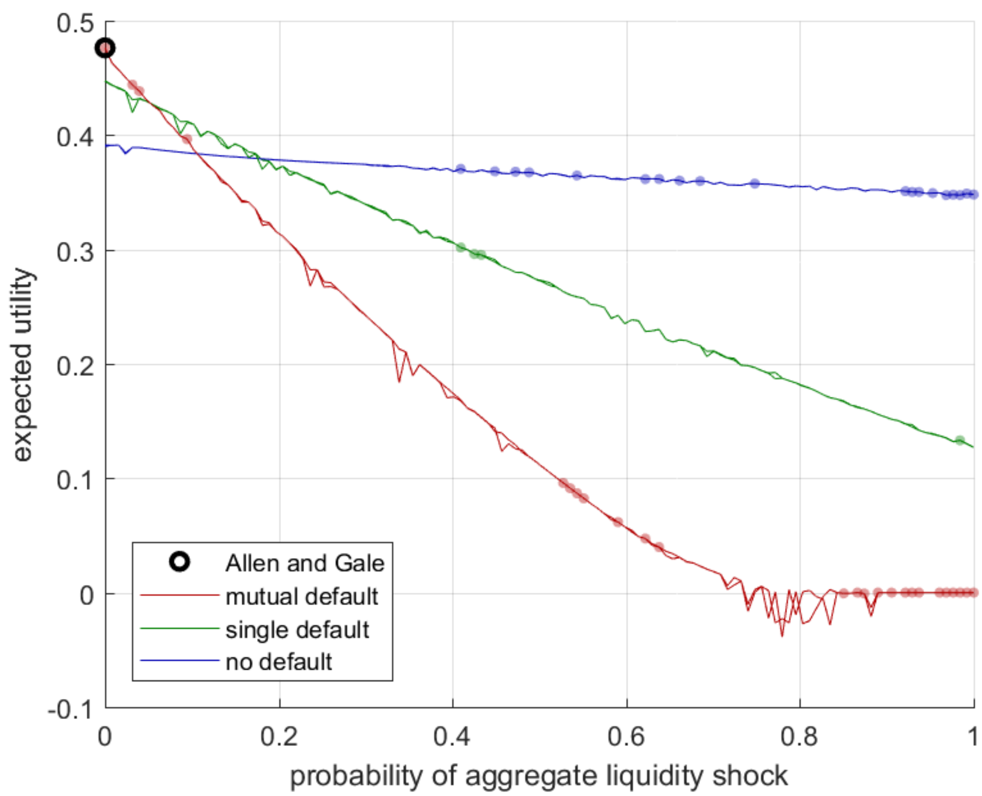
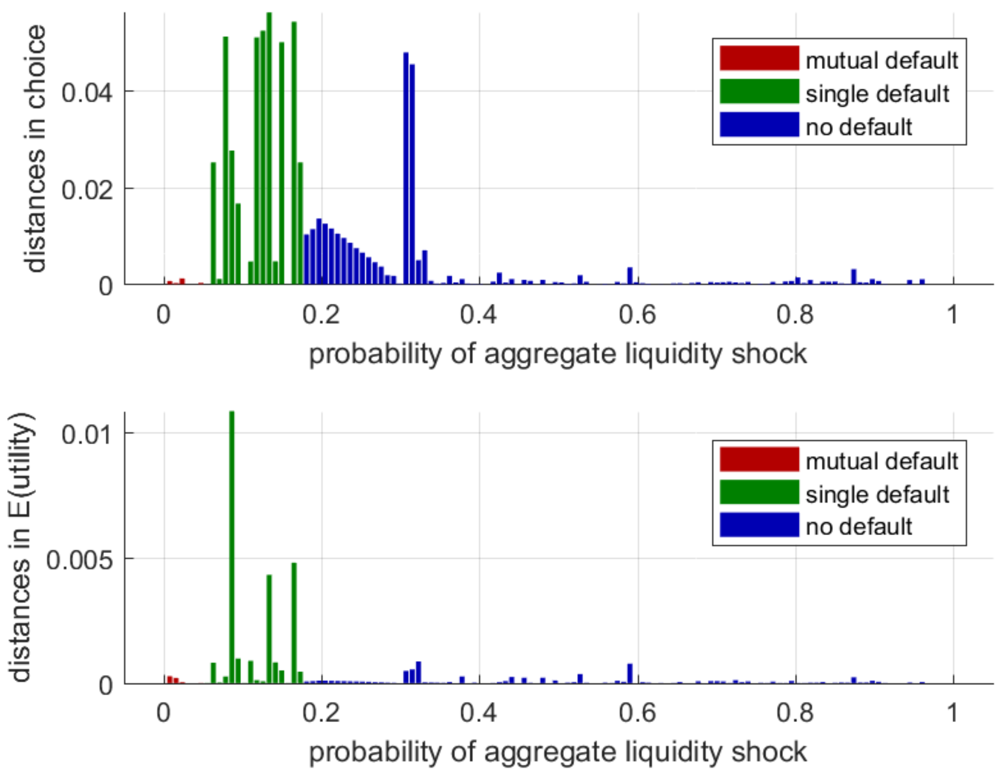


Figure 5.4: Numerical accuracy of the results presented in Figures 5.3 and 5.5 to 5.7. The top panel shows the ℓ_1 norm of the difference between the two choice vectors that constitute the computed equilibrium choice pair for each type of equilibrium, where each type is superior. The bottom panel shows the distance between the pair of expected utility values corresponding to the pair of equilibrium choice vectors. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$, $\alpha = 0.15$ and $\varepsilon = 0.1$.



I summarize the findings from these two parameter cases in the following result.

Numerical Result 2. *The type of Nash equilibrium depends on parameters.*

1. *There exist parameter regions (e.g. when α is sufficiently small) where the no-default equilibrium is weakly superior $\forall\, p \in [0,1]$; i.e. the mutual-default equilibrium (contagion) can only occur with probability* 0.

2. *There exist parameter regions (e.g. when α is sufficiently large) where the equilibrium type depends on the probability of the aggregate liquidity demand shock.*

   (i) *When the probability of the aggregate liquidity demand shock is low enough, the mutual-default equilibrium is superior to the single- and no-default equilibria.*

   (ii) *When the probability of the aggregate liquidity demand shock is intermediate, the single-default equilibrium is superior to the mutual- and no-default equilibria.*

   (iii) *When the probability of the aggregate liquidity demand shock is high enough, the no-default equilibrium is superior to the single- and mutual-default equilibria.*

The mutual-default equilibrium is also called the contagion case because the decision of late depositors in one region to run on their bank induces the late depositors in the other region to also run on their bank. I show that contagion occurs only for a low enough probability of the aggregate liquidity demand shock. For a higher shock probability, in the single-default equilibrium, banks internalize the aggregate liquidity risk to the extent that a run by late depositors of one bank no longer induces a run by late depositors of the other. For a high probability of the aggregate liquidity demand shock, banks fully internalize the ex-post risk of a shock ex ante and the equilibrium choice is such that default never occurs in equilibrium.

The intuition for the various types of equilibrium and their dependence on the probability of the aggregate liquidity demand shock is similar to the intuition for optimality of the two regimes, default and no default, of a global bank (chapter 3). This intuition is based on the ex-post inefficiency of satisfying the constraint of non-state-contingent deposit return across all states in which a bank does not default.

When the aggregate liquidity demand shock is large, a non-state-contingent deposit return in states requires a very low late consumption allocation in states 3 and 4 (where the shock realizes). This is ex-post inefficient. Thus, when the probability of the aggregate liquidity demand shock is very low (i.e. the weight on states 3 and 4 in ex-ante expected utility is low), it is efficient for a bank to choose an allocation where it defaults in the any state where the aggregate liquidity demand realizes. This allows the bank to choose a more efficient risk sharing allocation with deposit return that is equal across only in states 1 and 2. In other words: maintaining a non-state-contingent deposit return in all states, means that risk-sharing in all states is ex post inefficient. If the probability of the aggregate liquidity demand shock is low enough, it is ex-ante efficient to accept an *decrease* in ex-post efficiency in states 3 and 4 (by defaulting) in order attain an *increase* in ex-post efficiency in states 1 and 2. This is a situation where the Nash equilibrium is characterized by mutual default (contagion), and corresponds to the optimality of the default regime of a global bank. I show below that in this situation the banks choose mutual interbank positions that insure only against the regional liquidity demand shock: $z_A = z_B = \varepsilon$.

In contrast, when the probability of the aggregate liquidity demand shock is very high, the weight on states 3 and 4 in ex-ante expected utility is high, so default in either of these states is very costly relative to the cost of maintaining non-state-contingent deposit returns in all states. Thus both banks choose allocations such that they never default in any state. This is a situation where the Nash equilibrium is characterized by no default in any state, and corresponds to the optimality of the no default regime of a global bank. I show below that in this situation the banks choose mutual interbank positions that insure against the aggregate liquidity demand shock: $z_A = z_B = \alpha > \varepsilon$.

Finally, when the probability of the aggregate liquidity demand shock is intermediate, each bank prefers an allocation where it defaults only in the state where the aggregate liquidity demand shock hits its own region (state 4 for bank A and state 3 for bank B). This is a situation where the Nash equilibrium is characterized by a single default. Since a global bank is constrained to offer identical allocations in each region in any given state, there is no event in the benchmark allocation that corresponds to the single default equilibrium in the decentralized game. I show below that, in this situation, the banks choose mutual interbank positions that are smaller than the regional liquidity demand shock: $z_A = z_B < \varepsilon < \alpha$.

From here on, I focus on the parameter set where there are multiple equilibrium types across p (i.e. for large α), which is the more interesting case.

Figure 5.5 shows the equilibrium liquidity (y^*) across the probability of the aggregate liquidity demand shock (p) at my approximation of the symmetric Nash equilibrium for each type of equilibrium (no default, single default or mutual default) where each type is superior. The figure displays all of the iterates in the application of algorithm 4.4.3 for this parameter set, as well as the final equilibrium choice pair as a pair of lines.

There is clear evidence of a point of attraction which I interpret as a symmetric equilibrium. All iterates fall within a small sub-interval of the search space for this variable, $[0,1]$. The figure mirrors the results from Figure 5.4: the mutual-default equilibrium is most accurate, followed by the no-default equilibrium and then the single-default equilibrium.

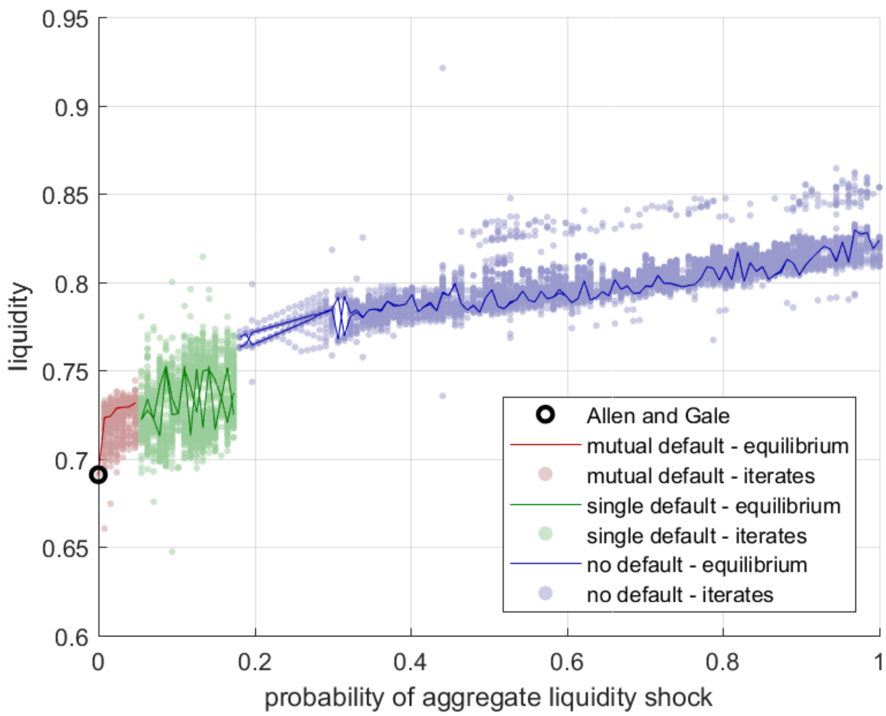
I summarize the findings from Figure 5.5 in the following result:

Numerical Result 3. *The symmetric Nash equilibrium liquidity (y^*) is increasing in the probability of the aggregate liquidity demand shock, with some evidence that it may be discontinuous across equilibrium types.*

(i) When the probability of the aggregate liquidity demand shock is low enough, the mutual-default equilibrium is superior. At $p = 0$, it is equal to the liquidity of the Allen and Gale (2000) allocation. As p increases, y^* increases until the equilibrium type switches to single default.

(ii) When the probability of the aggregate liquidity demand shock is intermediate, the single-default equilibrium is superior. There is suggestive evidence that y^* decreases discontinuously at the probability boundary with the mutual equilibrium type. As p increases, y^* is weakly increasing until the equilibrium type switches to no default.

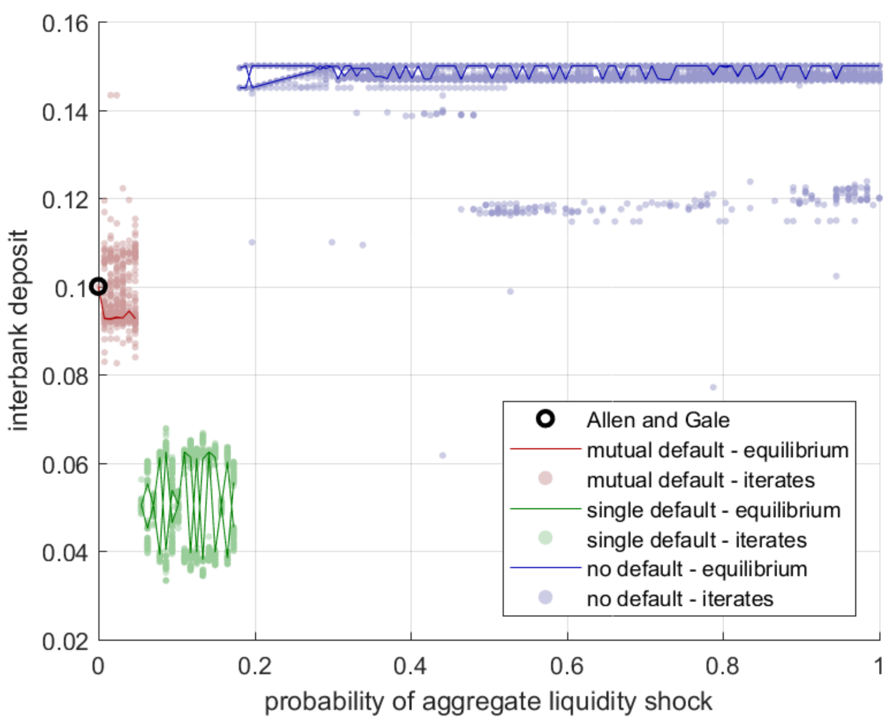
Figure 5.5: Liquidity at the approximate symmetric Nash equilibrium of the regional bank problem across the probability of the aggregate liquidity demand shock. The mutual-default equilibrium is superior for $p \in [0, 0.06)$, the single-default equilibrium for $p \in [0.06, 0.18]$, and the no-default equilibrium for $p \in (0.18, 1]$. When the probability is zero, the equilibrium liquidity is equal to that in the Allen and Gale (2000) allocation, within numerical precision. Equilibrium liquidity is generally increasing in p, but with discrete jumps between different equilibrium types. The search space is: $[0, 1]$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$, $\alpha = 0.15$ and $\varepsilon = 0.1$.



(iii) *When the probability of the aggregate liquidity demand shock is high enough, the no-default equilibrium is superior, and y^* increases discontinuously at the probability boundary with the single equilibrium type. y^* is increasing in p until $p = 1$.*

Figure 5.6 shows the equilibrium interbank deposit (z^*) across the probability of the aggregate liquidity demand shock (p) at my approximation of the symmetric Nash equilibrium for each type of equilibrium (no default, single default and mutual default) where each type is superior. As above, there is large variability in computations of the single-default equilibrium, where the mutual- and no-default equilibria are more tightly packed. Nevertheless, there is strong attraction to a symmetric equilibrium as all the iterates fall within a small sub-interval of the search space for this variable, $[0, 0.15]$.

Figure 5.6: Interbank deposit at the approximate symmetric Nash equilibrium of the regional bank problem across the probability of the aggregate liquidity demand shock. The equilibrium interbank deposit is approximately flat within each equilibrium type with discrete jumps between equilibrium types. When the probability is zero, the Allen and Gale (2000) allocation is replicated by the mutual-default equilibrium. When the probability is low enough ($p \in [0, 0.06)$), the mutual-default equilibrium is superior and the symmetric equilibrium interbank deposit is close to the regional liquidity shock ε. For intermediate probability values ($p \in [0.06, 0.18]$), the single-default equilibrium is superior and the equilibrium interbank deposit is far lower than ε. When the probability is high enough ($p \in (0.18, 1]$), the no-default equilibrium is superior and the equilibrium interbank deposit is approximately equal to the aggregate liquidity demand shock, α. The search space is $[0, 0.15]$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$, $\alpha = 0.15$ and $\varepsilon = 0.1$.



I summarize the findings from Figure 5.6 in the following numerical result:

Numerical Result 4. *The symmetric Nash equilibrium interbank deposit (z^*) is discontinuous across equilibrium types and there is a clear ranking in the size of the interbank deposit held in equilibrium. The highest position $z^* \approx \alpha$ is held in the no-default equilibrium, an intermediate level is held in the mutual-default equilibrium $z^* \approx \varepsilon$, and the lowest deposit is held in the single-default equilibrium $z^* < \varepsilon$.*

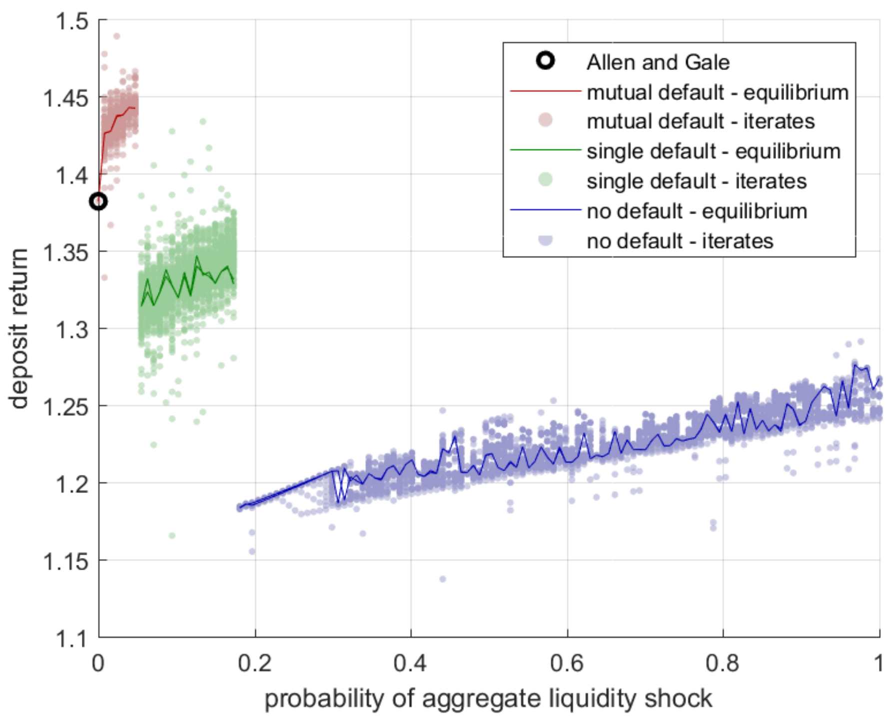
(i) When the probability of the aggregate liquidity demand shock is low enough, the mutual-default equilibrium is superior. At $p = 0$, it is equal to the interbank position of the Allen and Gale (2000) allocation. $z^* \approx \varepsilon$ and is non-increasing in p until the equilibrium type changes to single default.

(ii) When the probability of the aggregate liquidity demand shock is intermediate, the single-default equilibrium is superior. There is a discontinuous decrease in the interbank deposit to $z^* < \varepsilon$ at the probability boundary with the mutual-default equilibrium. There is no evidence of either an increasing or a decreasing tendency in p until the equilibrium type switches to no default.

(iii) When the probability of the aggregate liquidity demand shock is high enough, the no-default equilibrium is superior. There is a discontinuous increase in the interbank deposit to $z^* \approx \alpha$ at the probability boundary with the single-default equilibrium. z^* remains approximately equal to α until $p = 1$.

Figure 5.7 shows the equilibrium deposit return (d^*) across the probability of the aggregate liquidity demand shock (p) at my numerical approximation of the symmetric Nash equilibrium for each type of equilibrium where each type is superior. As in Figures 5.5 and 5.6, there is a clear attraction to a symmetric equilibrium as the iterates fall within a small sub-interval in the search space for this variable, $[1, 5]$.

Figure 5.7: Deposit return at the approximate symmetric Nash equilibrium of the regional bank problem across the probability of the aggregate liquidity demand shock. When the probability is zero, the Allen and Gale (2000) allocation is replicated by the mutual-default equilibrium. When the probability is low enough ($p \in [0, 0.06)$), the mutual-default equilibrium is superior. For intermediate probability values ($p \in [0.06, 0.18]$), the single-default equilibrium is superior and when the probability is high enough ($p \in (0.18, 1]$), the no-default equilibrium is superior. Within each equilibrium type the equilibrium deposit return is increasing, with clear downward jumps across equilibrium types, so that the mutual-default equilibrium has the highest deposit return, followed by the single- and then the no-default equilibria. The search space is $[1, 5]$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$, $\alpha = 0.15$ and $\varepsilon = 0.1$.



I summarize the findings from Figure 5.7 in the following numerical result:

Numerical Result 5. *The symmetric Nash equilibrium deposit return (d^*) is increasing in p within each equilibrium type and is discontinuous across equilibrium types. There is a clear ranking across equilibrium types: the deposit return is highest in the mutual-default equilibrium, followed by the single-default and then the no-default equilibria.*

(i) When the probability of the aggregate liquidity demand shock is low enough, the mutual-default equilibrium is superior. At $p = 0$, the numerical equilibrium attains the value in the Allen and Gale (2000) allocation. d^* is increasing in p until the equilibrium switches to single default.

(ii) When the probability of the aggregate liquidity demand shock is intermediate, the single-default equilibrium is superior. d^* decreases discontinuously at the probability boundary with the mutual-default equilibrium. d^* is increasing in p until the equilibrium type switches to no default.

(iii) When the probability of the aggregate liquidity demand shock is high enough, the no-default equilibrium is superior. d^* decreases discontinuously at the probability boundary with the single-default equilibrium. d^* is increasing in p until $p = 1$.

The discontinuities in the optimal choices obtain when the equilibrium switches from one type of dominant equilibrium (e.g. mutual-default) to another (e.g. single-default), mirroring the discontinuities in the benchmark allocations (chapter 3). These discontinuities are features of my solution rather than anomalies of the numerical approximation. First, the symmetric equilibrium choices of an arbitrary bank in each equilibrium have a continuous character (up to numerical precision). Second, the discontinuities between equilibrium types are much larger than any remaining numerical imprecision, which especially can be observed in Figures 5.6 and 5.7.

5.3 Results for the probability and size of the aggregate liquidity demand shock, accounting for risk aversion

In this section, I study the prevalence of contagion in terms of the two core parameters of this research: the probability p and size α of the aggregate liquidity demand shock. Additionally, I examine how the prevalence of contagion depends on the degree of risk aversion.

The main findings in this section are as follows. Contagion occurs when the size of the aggregate liquidity demand shock is large enough and its probability is small enough. The slope of the probability boundary between the no-default equilibrium and the other equilibrium types is positive in the size of the aggregate liquidity demand shock, mirroring the results for the global bank presented in Figure 3.3. Lastly, the size of the parameter region in α and p in which either the single- or mutual-default equilibria are superior is decreasing in the degree of risk aversion (ρ); that is, as consumers become more risk averse, contagion becomes less likely.

The results are presented as a sequence of figures across the full range of p and α for a range of increasing risk aversion parameters, ρ. In this section I consider a grid of values of p and α, concentrating on low values of p. That is, I have computed equilibrium types for a 16×16 grid with $p \in [0, 0.1]$ and $\alpha \in [0, \frac{1-\gamma}{2}]$, and for a second 16×16 grid with $p \in [0.1, 1]$ and $\alpha \in [0, \frac{1-\gamma}{2}]$, yielding 512 parameter sets per risk aversion parameter. Table 5.5 summarizes the model parameters studied in this section.

Table 5.5: Model parameters for the results across the probability and size of the aggregate liquidity demand shock.

Parameter Name	Symbol	Value
Utility function	$u(c)$	$\frac{c^{1-\rho}-1}{1-\rho}$
Coefficient of risk aversion	ρ	$\{2,3,\ldots,6\}$
Investment return at maturity	R	5
Investment return at early liquidation	r	0.1
Average liquidity demand	γ	0.5
Regional liquidity demand shock	ε	0.1
Aggregate liquidity demand shock size	α	$[0,0.25]$
Aggregate liquidity demand shock probability	p	$[0,1]$

I summarize the findings from Figures 5.8 to 5.12 in the following numerical result which corresponds to the benchmark result, presented in Figure 3.3:

Numerical Result 6.

(i) Contagion occurs when the probability of the aggregate liquidity demand shock p is low enough and the size α of the shock is large enough.

(ii) The parameter space in the probability p and size α of the aggregate liquidity demand shock where contagion occurs is decreasing in the degree of risk aversion (ρ).

(iii) The boundary $\underline{p}(\alpha)$ above which no default occurs is increasing in α.

Figure 5.8: Equilibrium types at the approximate symmetric Nash equilibrium of the regional bank problem across the probability p and size α of the aggregate liquidity demand shock for risk aversion parameter $\rho = 2$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $R = 5$, $r = 0.1$, $\gamma = 0.5$ and $\varepsilon = 0.1$.

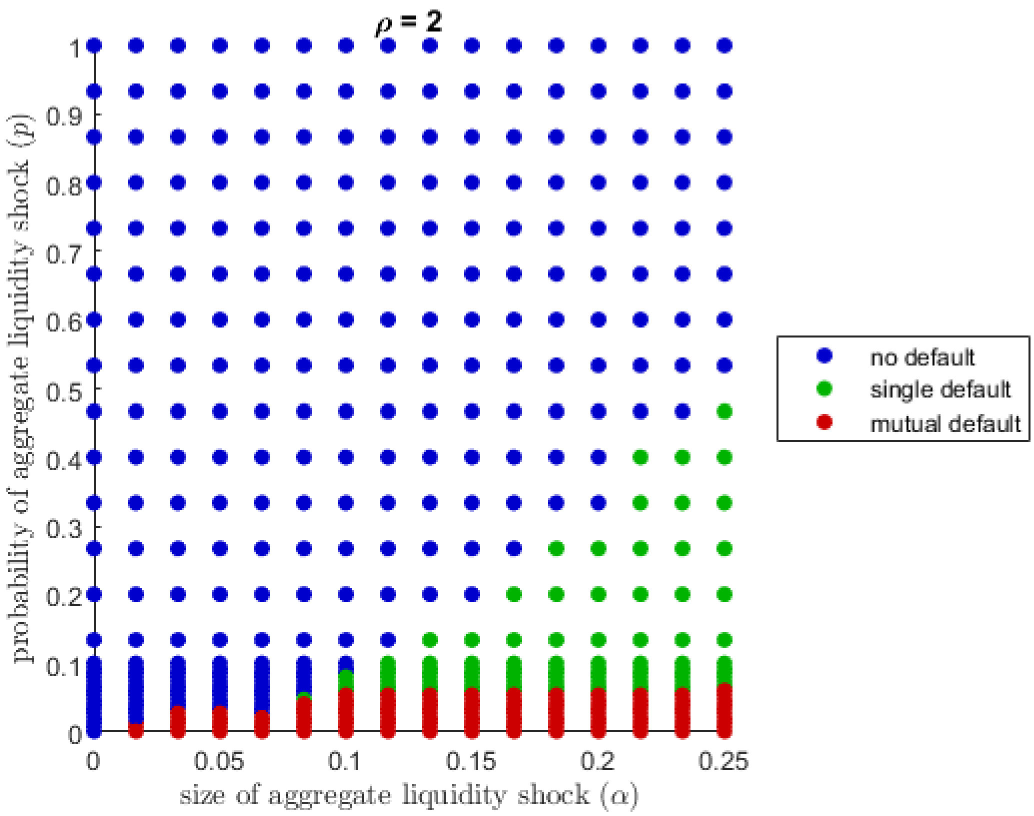


Figure 5.9: Equilibrium types at the approximate symmetric Nash equilibrium of the regional bank problem across the probability p and size α of the aggregate liquidity demand shock for risk aversion parameter $\rho = 3$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $R = 5$, $r = 0.1$, $\gamma = 0.5$ and $\varepsilon = 0.1$.

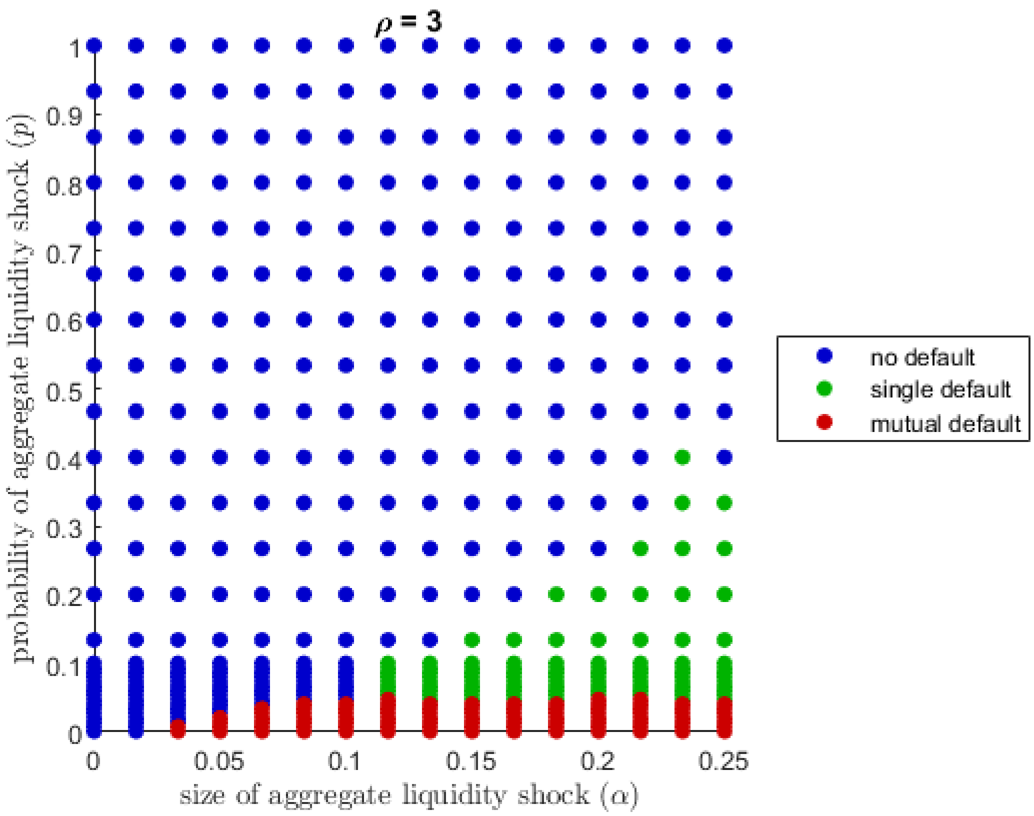


Figure 5.10: Equilibrium types at the approximate symmetric Nash equilibrium of the regional bank problem across the probability p and size α of the aggregate liquidity demand shock for risk aversion parameter $\rho = 4$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $R = 5$, $r = 0.1$, $\gamma = 0.5$ and $\varepsilon = 0.1$.

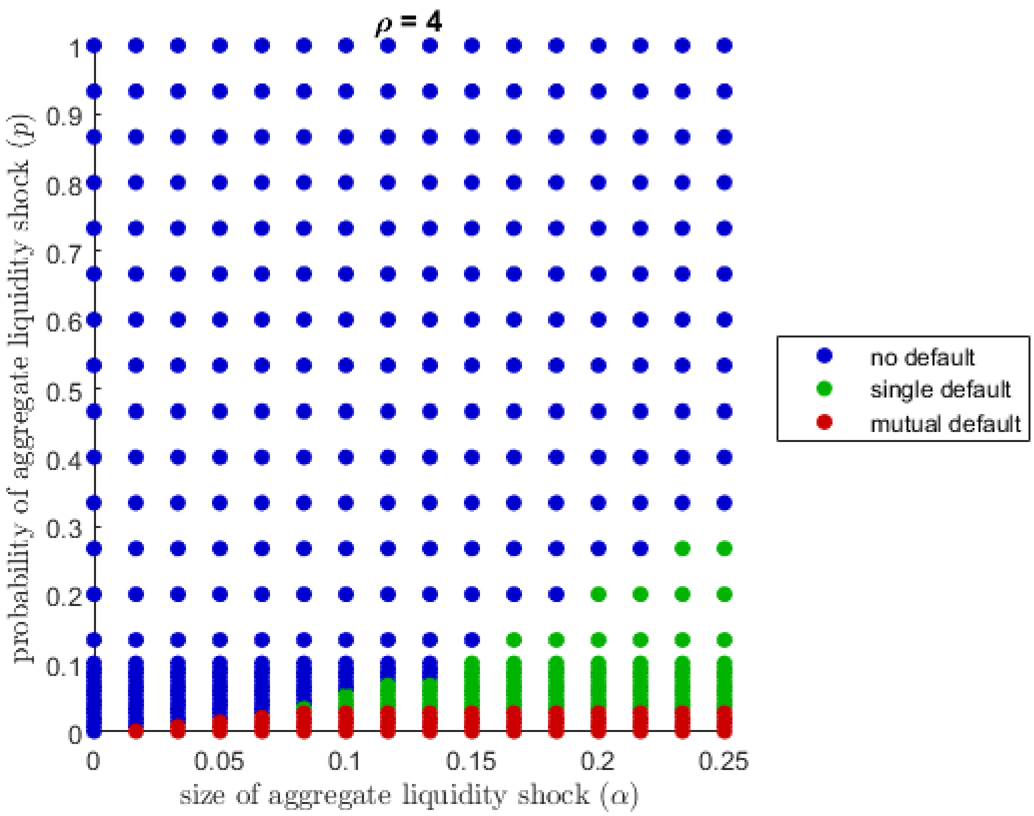


Figure 5.11: Equilibrium types at the approximate symmetric Nash equilibrium of the regional bank problem across the probability p and size α of the aggregate liquidity demand shock for risk aversion parameter $\rho = 5$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$ $R = 5,\ r = 0.1,$ $\gamma = 0.5$ and $\varepsilon = 0.1$.

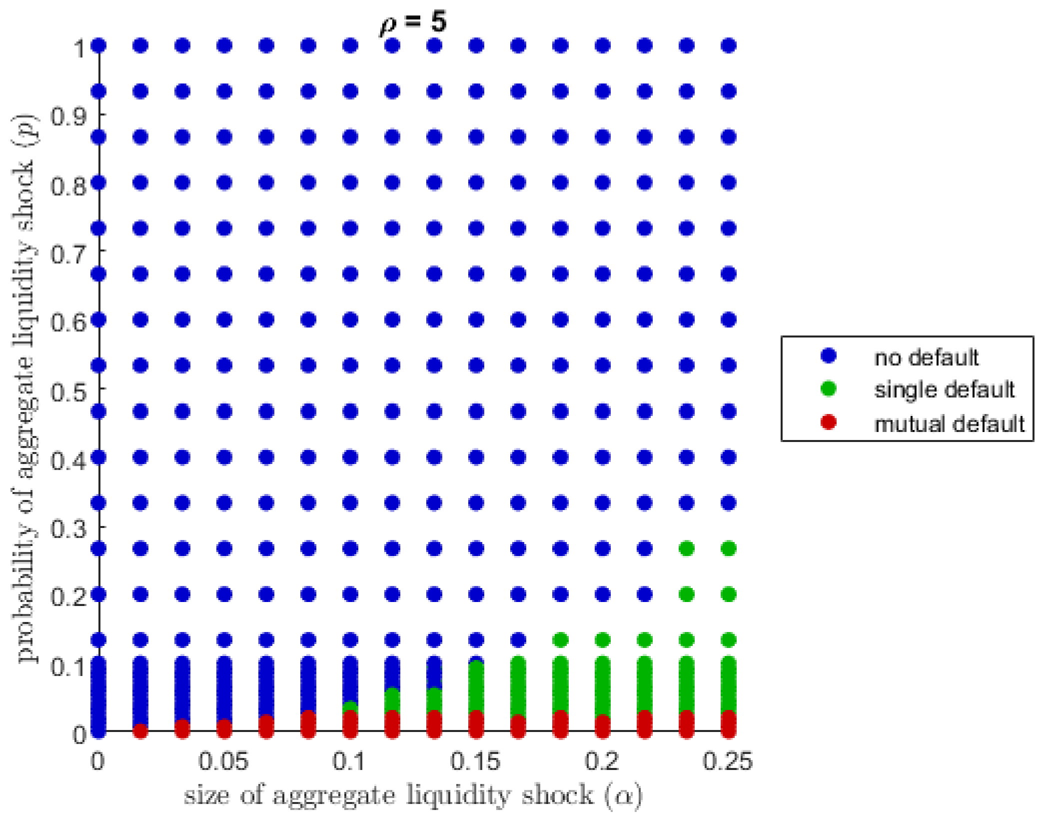
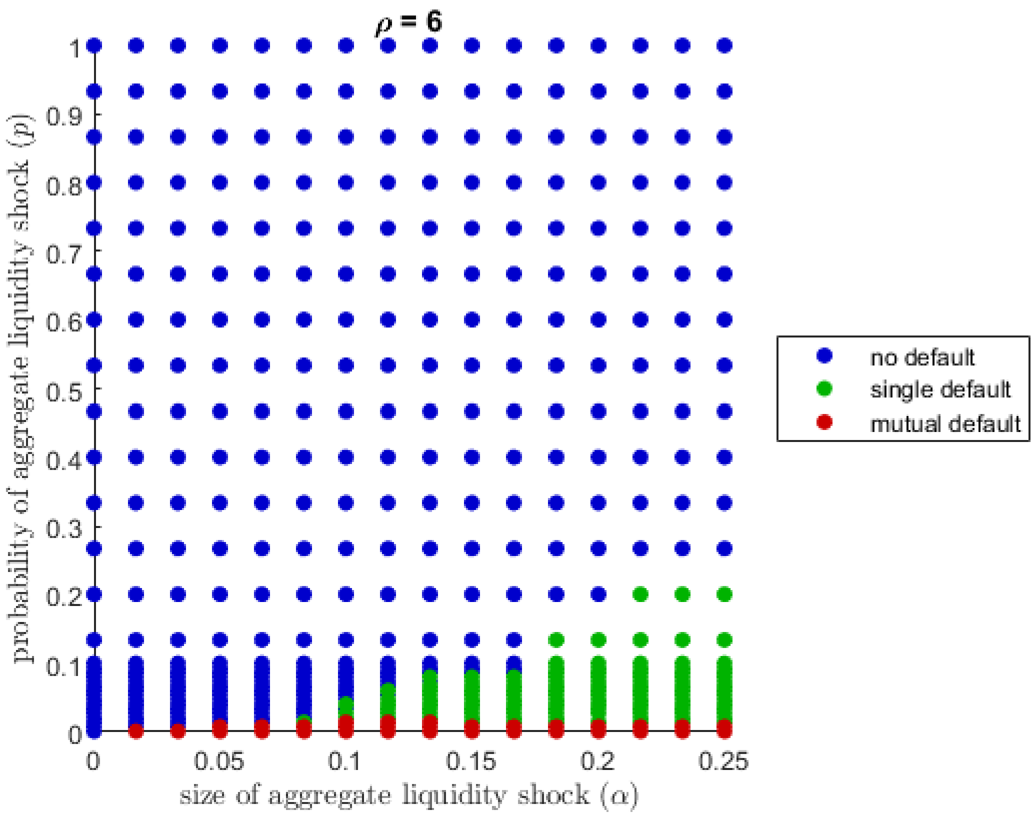


Figure 5.12: Equilibrium types at the approximate symmetric Nash equilibrium of the regional bank problem across the probability p and size α of the aggregate liquidity demand shock for risk aversion parameter $\rho = 6$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$ $R = 5$, $r = 0.1$, $\gamma = 0.5$ and $\varepsilon = 0.1$.



5.4 Prevalence of contagion across the full parameter space

In order to obtain general results that mimic those found using standard analytical methods, the entire parameter space must be characterized. Thus, in this section I present results on the prevalence of contagion across random draws from the full parameter space.

The main finding in this section is that contagion is rare. It is an equilibrium of the decentralized problem for only approximately 4% of the parameter space. This is an upper bound on the prevalence of contagion. Even in a parameter set where contagion is possible, it only occurs if the aggregate liquidity demand shock is realized, which happens with probability p. Therefore, I constructed a measure of contagion where its possibility (a parameter set where the mutual-default equilibrium is superior) is weighted by the probability of the aggregate liquidity demand shock. This measure gives an ex-ante likelihood for an arbitrary parameter draw of approximately 0.5%.

Table 5.6 presents the parameter space from which 6781 draws were taken. For each integer value of the risk aversion parameter $\rho \in \{2, 3, ..., 7\}$, the other parameters were drawn from independent, uniform distributions[1].

I present results on the prevalence of contagion using three metrics. The first is a direct numerical measure of the parameter space in which contagion can occur, denoted $\overline{\mu}_{contagion}$. This metric is simply the frequency of parameter draws that are such that the mutual-default equilibrium is superior.

The second metric is a weighted numerical measure of the parameter space in which contagion occurs. The weight on each occurrence of a parameter set where the mutual-default equilibrium is superior is the probability of the aggregate liquidity demand shock in that parameter set. The second metric is denoted $\underline{\mu}_{contagion}$. The third metric, denoted $\mu_{contagion}$, is an intermediate version of the previous two, where probability-weighted measure is standardized

[1]Since the components of the early liquidity demand specification must sum to less than or equal to one, the admissible ranges of the regional (ε) and aggregate (α) liquidity demand shocks depend on the draw of the average liquidity demand in the absence of an aggregate shock (γ). Thus, the independence of the draws are conditional on their admissible range.

Table 5.6: Model parameters for results across the full range of parameters

Parameter Name	Symbol	Value/Range
Utility function	$u(c)$	$\frac{c^{1-\rho}-1}{1-\rho}$
Risk aversion	ρ	$\{2,3,...,7\}$
Aggregate liquidity demand shock probability	p	$(0,1)$
Investment return at maturity	R	$(1,10)$
Investment return at early liquidation	r	$(0,1)$
Average liquidity demand	γ	$(0,1)$
Aggregate liquidity demand shock size	α	$(0,\frac{1-\gamma}{2})$
Regional liquidity demand shock size	ε	$(0,\min\{\alpha,\gamma\})$
Total number of draws	N	6784

by the sum of probabilities, rather than by a count of all parameter sets (see the mathematical definitions of the metrics below).

The first metric of the prevalence of contagion is constructed as follows. For every parameter draw i, the equilibrium type is computed. I construct an indicator variable $i_{contagion}$, where $i_{contagion} = 1$ if the equilibrium type is mutual default (contagion). For the no- or single-default equilibrium, $i_{contagion} = 0$. The metric is then simply the sum of the indicator variable standardized by the total number of parameter sets considered:

$$\overline{\mu}_{contagion} = \frac{1}{N}\sum_{i=1}^{N} i_{contagion}.$$

This first metric is an upper bound on the likelihood of contagion. Even within a parameter draw i where the mutual-default equilibrium is superior, contagion occurs only if the aggregate liquidity demand shock is realized, which happens with the probability of the aggregate liquidity demand shock in that particular draw, p_i. Thus, the second metric, a lower bound on the prevalence of contagion, is calculated as the sum of probabilities in the parameter sets where contagion is possible, standardized by the total number of parameter sets considered:

$$\underline{\mu}_{contagion} = \frac{1}{N}\sum_{i=1}^{N} p_i \cdot i_{contagion}.$$

The second metric is a lower bound on the likelihood of contagion. The sum of probabilities is divided by the number of parameter sets. However, this may be too strict. As such, I consider an intermediate metric which standardizes the sum of probabilities from parameter sets where contagion occurs by the sum of probabilities in all parameter sets:

$$\mu_{contagion} = \frac{\sum_{i=1}^{N} p_i \cdot i_{contagion}}{\sum_{i=1}^{N} p_i}.$$

Table 5.7: The prevalence of contagion across the full parameter space

	Risk aversion parameter ρ						
	All	2	3	4	5	6	7
N	6784	3206	713	777	715	966	333
$\overline{\mu}_{contagion}$	0.037	0.039	0.039	0.048	0.029	0.027	0.033
$\mu_{contagion}$	0.005	0.004	0.007	0.007	0.005	0.005	0.009
$\underline{\mu}_{contagion}$	0.003	0.002	0.003	0.003	0.002	0.003	0.004

Table 5.7 presents the results on the prevalence of contagion based on the three metrics described above. I summarize the findings in the following numerical result.

Numerical Result 7. *Contagion is rare, occurring in approximately only* 4% *of the parameter space of this model. The ex-ante probability of contagion for an arbitrary point in the parameter space lies between* 0.3% *and* 0.5%.

5.5 Equilibrium type classification by machine learning

For the final section of results on the strategic bank problem, I collated all computations and used machine learning to characterize the results over the model parameters. Fundamentally, I faced a classification problem: determining which parameter sets yield each of the equilibrium types, including contagion.

I used machine learning to disentangle the effects of the seven model parameters generally: in sections 5.2 and 5.3 I show results in one or a pair of parameters for carefully selected set of fixed values of the other parameters. The selection of fixed parameters was done to obtain regions in the varying parameters where each of the equilibrium types has positive measure. For a fully general characterization, I needed a robust method of constructing the mapping from parameter set to equilibrium type for the full ranges of all parameters. Machine learning is ideally suited to this multi-dimensional classification problem.

I used support-vector machines (SVMs) to classify equilibrium types as functions of the parameters of the model. To this end I provide a simplified review of the method applied to a case where there are two classes that can be fully separated by two features. Extensions of these algorithms have been developed to construct the optimal boundaries between multiple categories that cannot be fully separated (Boser et al., 1992; Cortes and Vapnik, 1995; Fung and Mangasarian, 2005).

5.5.1 Simulated example: perfect classification by linear support-vector machine

Consider an artificial example where some variable, c, falls into one of two categories, $c = \{-1, 1\}$, as a function of two features, $f_1, f_2 \in (0,1)$. Specifically let:

$$c = \begin{cases} -1 & \text{if } f_1 < 1 - f_2 \\ 1 & \text{if } f_1 \geq 1 - f_2 \end{cases}$$

I generated 100 draws from this process, with $f_{1,i}, f_{2,i} \underset{i.i.d}{\sim} \text{uniform}(0,1)$ and fitted a linear SVM to classify the output. The set of observations $\{f_{1,i}, f_{2,i}, c_i\}_{i=1}^{100}$ is called the *training sample* of *patterns*. A linear SVM is an algorithm that finds a separating hyperplane $D : [0,1] \times [0,1] \to \mathbb{R}$ such that $D(f_1, f_2) < 0$ implies $c(f_1, f_2) = -1$, and $D(f_1, f_2) > 0$ implies $c(f_1, f_2) = 1$. The hyperplane $D(f_1, f_2)$ is called the *decision function* and is defined by a subset of patterns in the training sample, called *support vectors*. The locus defined by $D(f_1, f_2) = 0$ is called the *decision boundary*, and the algorithm is built to maximize the margin, which is defined the distance between the decision boundary and the closest elements of the set of $f_{1,i}, f_{2,i}$ pairs in either category. The closest elements of the set of $f_{1,i}, f_{2,i}$ pairs in either category are the vectors that support the separating hyperplane (decision function). The estimated SVM is used predictively by inputting new feature data and predicting the classification by means of the estimated decision function

Figures 5.13 and 5.14 illustrate the results of this simple example of fitting an SVM to perfectly separable data. Figure 5.13 shows the results in two dimensions. The figure displays the different categories of outcomes (c), and how these depend on the features (f_1, f_2), the decision boundary ($D(f_1, f_2) = 0$) and the support vectors that define the separating hyperplane (decision function, $D(f_1, f_2)$). Figure 5.14 presents the same example in three dimensions, to illustrate the graph of the decision function $D(f_1, f_2)$ as a hyperplane in $\mathbb{R}^3$.

5.5.2 Results for classification of equilibrium type by support-vector machines

In this section I show the results of employing SVMs on my computational results to gain insight into the values of the parameters that yield different equilibrium outcomes. The categories of outcomes I wish to classify are the three equilibrium types (no default, single default and mutual default). The features used for the classification are the seven model parameters (see Table 5.6). For this I need a more general version of an SVM algorithm than was presented in section 5.5.1:

Figure 5.13: A simple example of classification by a linear support-vector machine in two dimensions.

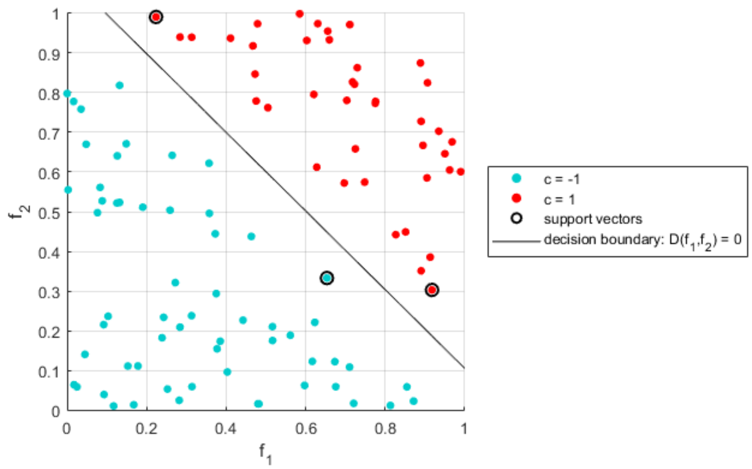
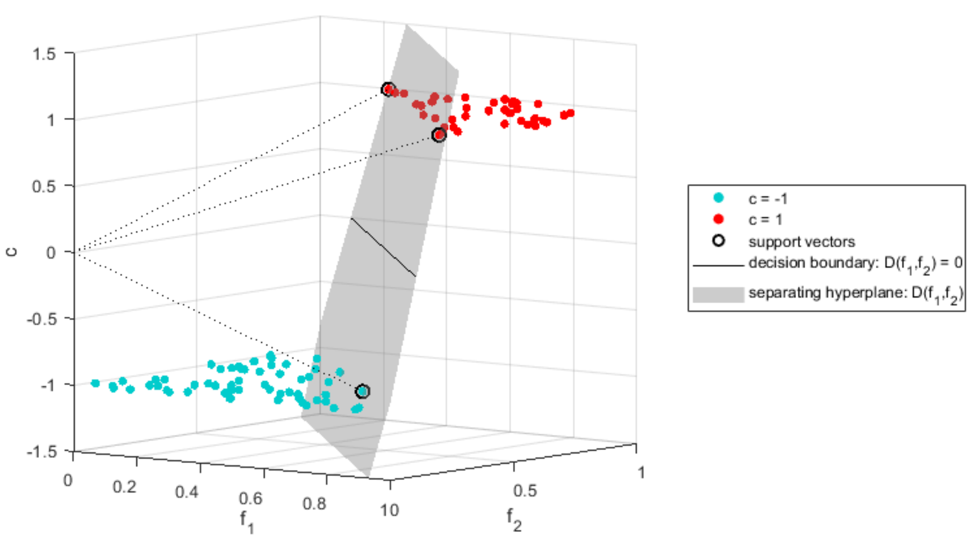


Figure 5.14: A simple example of classification by a linear support-vector machine in three dimensions.



(i) There are three equilibrium types, therefore a multi-category extension of the simple two category SVM was required,

(ii) Due to numerical error, not all equilibrium types were accurately computed; hence an SVM algorithm robust to error (also called a "soft-margin" SVM) was necessary, and

(iii) It is not obvious that the parameter boundaries between equilibrium types are best approximated by linear functions, hence a non-linear extension of the SVM algorithm was used.

I used the machine learning algorithms provided in Matlab (2017), and applied them to all the computational results I generated for all of the sub-questions studied in sections 5.1 to 5.4. The overall accuracy of three SVM specifications trained on a total of 15 552 computed patterns is presented in Table 5.8.

Table 5.8: Accuracy of various SVM algorithms in classifying equilibrium type as a function of model parameters

Type of decision function	Accuracy
Linear	91.8% equilibrium types accurately classified
Quadratic	95% equilibrium types accurately classified
Cubic	93.6% equilibrium types accurately classified

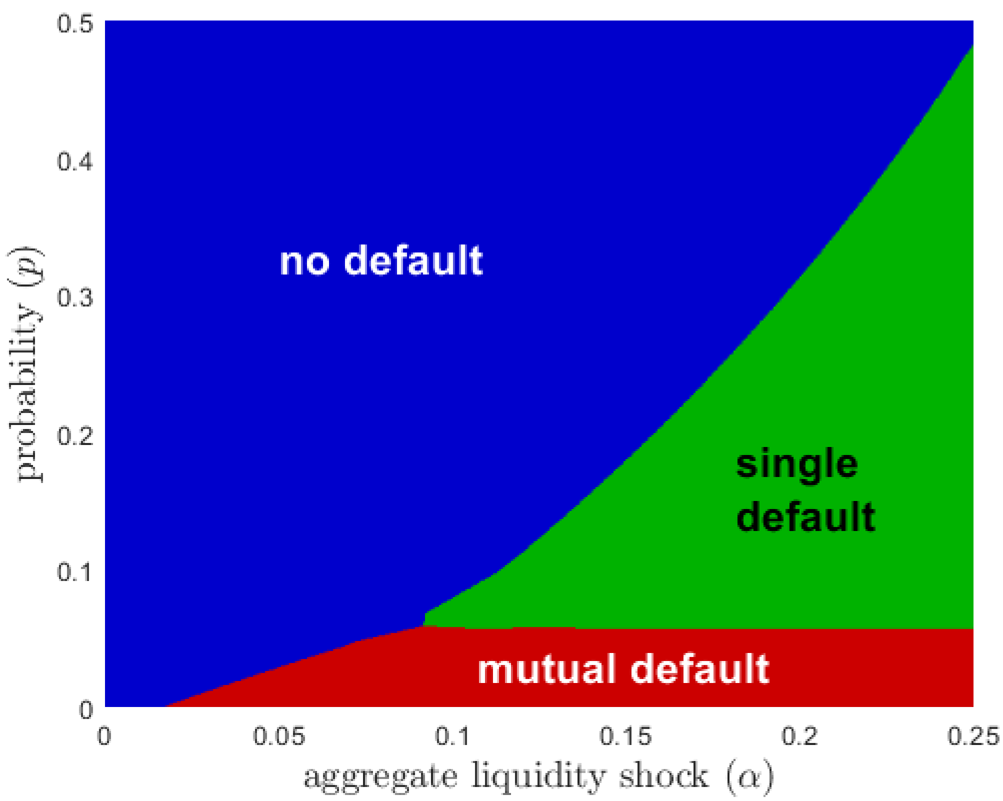
I present my findings from training SVMs in Figures 5.15 to 5.17. Since the quadratic SVM provided the highest accuracy (see Table 5.8), I present all results based on this SVM. Note that in each case, a set of fixed parameters was chosen to maximize the relative size of all equilibrium types, in order to build intuition regarding the impact of each parameter for the nature of the equilibrium type. In each figure, the boundary between equilibrium types corresponds to the non-linear, multi-category extension of the decision boundary between categories presented in section 5.5.1. The figures for the pairs of parameters are generated by simulating the predicted equilibrium type, based on the fitted SVM on a 500×500 regularly spaced grid for each of the parameter pairs.

Figure 5.15 presents the SVM-simulated equilibrium types across the size (α) and probability (p) of the aggregate liquidity demand shock. These results are essentially a replication of the grid based results presented in section 5.3, except they are based on all computations, rather than just on a grid in the two parameters with the others fixed. Despite the greater generality, the results remain the same. Contagion occurs only for low enough probability and large enough size of the aggregate liquidity demand shock. The single-default equilibrium is only optimal for an intermediate probability and large enough size of the aggregate liquidity demand shock. The probability boundary between the no-default equilibrium and the single- and mutual-default equilibria is increasing in the size of the shock. There is no evidence of a non-zero slope of the probability boundary in α between the mutual default and single-default equilibrium types.

Figure 5.16 presents the SVM-simulated equilibrium types across the early liquidation return (r) and the return at maturity (R) of investment. For these parameters, I do not have grid based equivalents as I have for the results presented above. As such, the boundaries between equilibrium types are likely to be more imprecise, so I do not attempt to give a very strong interpretation to their slopes where they are irregular.

I restrict my interpretation of these SVM results to the prevalence of contagion in the parameters r and R. When $r = 1$, liquidity is redundant, as there is no penalty to early liquidation of the investment, thus contagion never occurs. Contagion occurs when the early liquidation return is intermediate and the return at maturity of the investment is large enough, with the lower bound on R increasing in r. The intuition for this is as follows. When the return at maturity of investment is low, and the early liquidation return is high, investment is closer to liquidity than otherwise. As such, it is optimal to choose high liquidity holdings, which makes the no-default equilibrium more likely to be an equilibrium. However, when maturity returns are high and early returns are intermediate, investment is clearly dominant for financing greater late consumption. Then it becomes an equilibrium to risk contagion, as a large investment is de-

Figure 5.15: Simulated equilibrium types across the size, α, and probability, p, of the aggregate liquidity demand shock based on a fitted quadratic support-vector machine. The upper bound on p was chosen to focus attention on the region of interest, whereas the upper bound on α is due to the restriction that $\alpha \leq \frac{1-\gamma}{2}$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$ and $\varepsilon = 0.1$.



sirable. Lastly, when the early liquidation return is very low but the return at maturity is high enough, the cost of default becomes very high, and banks choose allocations that do not allow contagion.

Figure 5.17 presents the SVM-simulated equilibrium types across the average liquidity demand in the absence of an aggregate liquidity demand shock (γ) and the size of the regional liquidity demand shock (ε). Contagion is a Nash equilibrium when both γ and ε are high enough. First, the largest upper bound on the regional liquidity demand shock is obtained when $\gamma = 0.5$, since the probability features of the model require that $\varepsilon \leq \min\{\gamma, 1-\gamma\}$. Second, when ε is high, there is large variation in the number of early consumers across states in the absence of an aggregate liquidity demand shock, which creates a large incentive for mutual liquidity insurance. This can induce contagion when the aggregate liquidity demand shock is realized (Note that for this example the probability of the aggregate liquidity demand shock is low: $p = 0.02$). By contrast, when the regional liquidity demand shock is small, the exposure of one bank to the default of another is similarly small, and the default of one bank does not tend to cause contagion.

As a final exercise, I repeated the above simulations for a set of increasing risk aversion parameters, to evaluate the impact of risk aversion on the prevalence of contagion as in section 5.3, but for all the parameters pairs considered in this section. The results are presented in Table 5.9 as the fraction of values in the 500×500 grid where the predicted equilibrium type from the fitted SVM is mutual default. For comparability with the figures, for each parameter pair in the grid, the fixed parameters are the same as those of the corresponding figure (Figures 5.15, 5.16 and 5.17), except for the risk aversion parameter. That is, the values in the table are the relative size of the mutual-default regions in the equivalents of the figures above.

Figure 5.16: Simulated equilibrium types across the early liquidation return, r, and the return at maturity, R, of investment based on a fitted quadratic support-vector machine. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $p = 0.09$, $\alpha = 0.1$, $\gamma = 0.6$ and $\varepsilon = 0.1$.

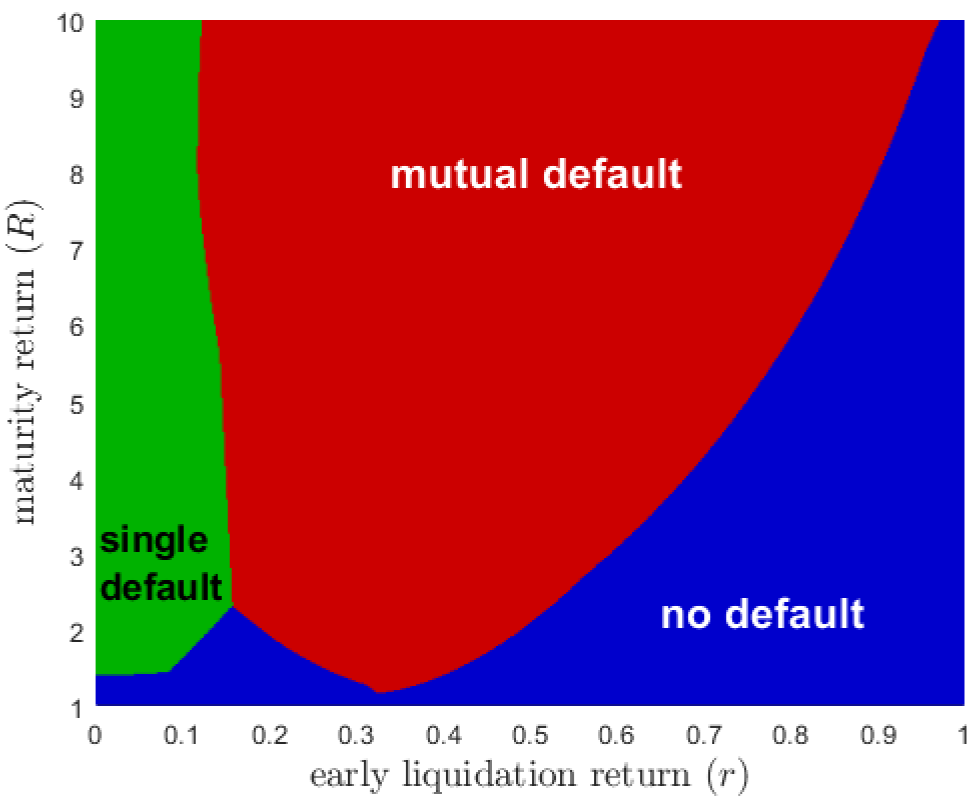


Figure 5.17: Simulated equilibrium types across the average liquidity demand in the absence of an aggregate liquidity demand shock, γ, and the size of the regional liquidity demand shock, ε, based on a fitted quadratic support-vector machine. The angled lower bound of the graph is due to the restriction that $\varepsilon \leq \gamma$, and the upper bound on γ is due the restriction that $\gamma \leq 1-2\alpha$. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $p = 0.02$, $\alpha = 0.1$, $R = 2$ and $r = 0.9$.

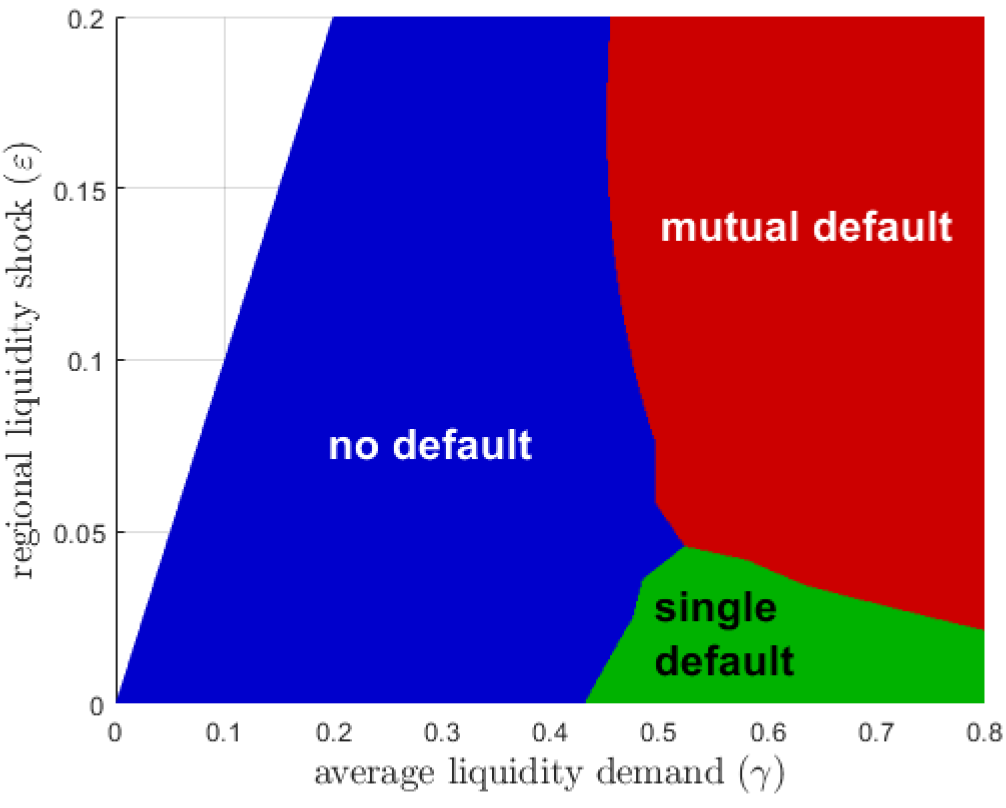


Table 5.9: Fraction of contagion outcomes in pair-wise parameter sets over increasing degrees of risk aversion.

	Risk aversion parameter ρ					
	2	3	4	5	6	7
Contagion fraction across α and p ($R = 5$, $r = 0.1$, $\gamma = 0.5$ and $\varepsilon = 0.1$)	0.092	0.070	0.051	0.034	0.018	0.005
Contagion fraction across r and R ($p = 0.09$, $\alpha = 0.1$, $\gamma = 0.6$ and $\varepsilon = 0.1$)	0.590	0.463	0.355	0.244	0.121	0.000
Contagion fraction across γ and ε ($p = 0.02$, $\alpha = 0.1$, $R = 2$ and $r = 0.9$)	0.293	0.307	0.323	0.340	0.360	0.383

As in section 5.3, the SVM simulations confirm that the fraction of the parameter space in the probability (p) and size (α) of the aggregate liquidity demand shock in which contagion occurs is decreasing in the degree of risk aversion. The same is true for the contagion fraction of the parameter space in the early liquidation return (r) and the return at maturity of investment (R).

In contrast, the fraction of the parameter space in the average liquidity demand in the absence of an aggregate liquidity demand shock (γ) and the regional liquidity demand shock (ε) in which contagion occurs is increasing in the degree of risk aversion. This can be explained as follows. The regional liquidity demand shock provides the incentive for liquidity insurance in the absence of an aggregate liquidity demand shock. As the degree of risk aversion increases, this insurance incentive increases; thus, exposure to contagion becomes an equilibrium for a larger parameter space in the regional liquidity demand shock.

I summarize the findings in this section in the following numerical result:

Numerical Result 8. *Contagion occurs when:*

(i) the aggregate liquidity demand shock has small enough probability (p) and large enough size (α),

(ii) investment has an intermediate early liquidation return (r), and the return at maturity (R) is large enough relative to the early liquidation return, and

(iii) the average liquidity demand in the absence of aggregate risk (γ) and the regional liquidity demand shock (ε) are large enough.

Additionally, the fraction of the pair-wise parameter space in which contagion occurs when the aggregate liquidity demand shock is realized is

(i) decreasing in risk aversion for the probability (p) and size (α) of the aggregate liquidity demand shock, as well as for the early liquidation return (r) and the return at maturity (R) of investment, but

(ii) increasing in risk aversion for the average liquidity demand (γ) and the size of the regional liquidity demand shock (ε).

Chapter 6

Welfare Analysis

In this section I compare the welfare implications of the numerical results in chapter 5 with the analytic benchmarks in chapter 3. There are two key findings. First, the mutual-default and no-default equilibria achieve very close to the expected utility of the global bank allocation (with full or asymmetric information), with differences that are small enough to be considered artefacts of numerical imprecision. Second, the single-default equilibrium often achieves expected utility greater than that of the global bank allocation, and these differences are too large to be due to numerical imprecision. The second finding has the following interpretation: an economy with a global bank which operates regional branches can be sub-optimal, as it disallows default in only one region at a time[1].

I proceed in two steps. First, I present, in graphical form, a finding across only the probability of the aggregate liquidity demand shock in order to highlight the main issues, comparable to the numerical results in section 5.2. Then, to cover the full parameter space, I use all computational results. I use both graphical and numerical metrics to provide the comparison between the benchmarks and the decentralized problem.

[1]This result is tangentially similar to the result in Castiglionesi et al. (2017) who consider a setting without liquidation where integration of regional banks is not always optimal. As banks become more integrated, the returns that their consumers at different banks receive more correlated returns. In the model presented here, a single banking entity with regional branches can be roughly seen as the limit of integration where the regional branches give identical allocations irrespective of region. Wagner (2010) presents a model where banks can diversify into similar projects, and there as well, full diversification makes banks identical.

Figure 6.1 shows the expected utility obtained in the three equilibrium types across the probability of the aggregate liquidity demand shock (p) in comparison with the expected utility in the global bank allocation under asymmetric information[2]. The expected utility of the global bank allocation is marginally larger than the expected utility in the mutual-default equilibrium, where the mutual-default equilibrium is superior to the no-default equilibrium. Similarly, the expected utility of the global bank is marginally larger than the expected utility in the no-default equilibrium, where the no-default equilibrium is superior to the mutual-default equilibrium. However, in the region where the single-default equilibrium is superior, the global bank expected utility is much lower than that obtained in the strategic bank problem. Thus, in this parameter region, a global bank with regional branches is inferior as it does not allow default in only one region.

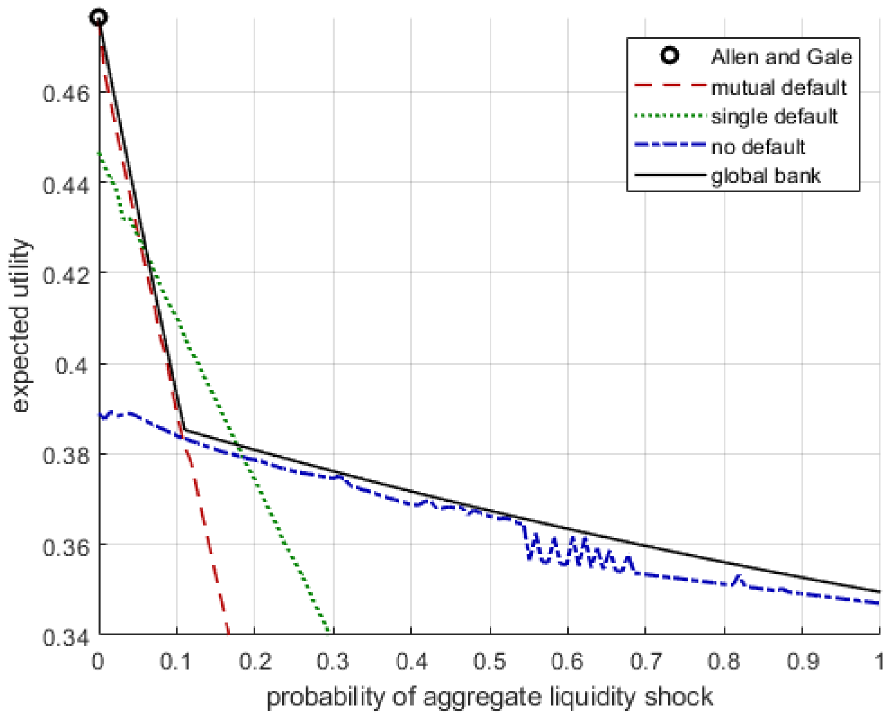
In Figures 6.2 and 6.3, and in Table 6.1 I consider results from 10207 unique parameter vectors from all computational results (including the random draws from the full parameter space from section 5.4).

Figure 6.2 presents scatter plots of the expected utility of the global bank allocation (under asymmetric information) and the expected utility of the three Nash equilibrium types of the strategic problem against the expected utility of the global bank allocation under full information.

In section 3.2 I show that there are parameter regions where the global bank allocation under asymmetric information differs from the allocation under full information, and therefore must yield lower expected utility. The top-left panel of Figure 6.2 shows that in the parameters covered by my computational results these cases are rare and that the expected utility differences are minor. The reason for the rarity and limited impact is as follows. First, when the default regime is optimal for the global bank under full information, it is also optimal under asymmetric information, as it is incentive compatible by construction. The default regime is optimal when the aggregate liquidity demand shock is large but has low probability (see Figure 3.3). The region where the global bank allocation under full information is not incentive com-

[2]For these parameters, the global bank allocations under full or asymmetric information are numerically identical.

Figure 6.1: Expected utility comparison between that attained by the global bank and at the approximate symmetric Nash equilibrium of the regional bank problem across the probability of the aggregate liquidity demand shock. The expected utility in the global bank allocation is marginally larger than in the mutual-default equilibrium (where the mutual-default equilibrium is superior to no-default equilibrium) and in the no-default equilibrium (where the no-default equilibrium is superior to the mutual-default equilibrium). However, in the region where the single-default equilibrium is superior, the global bank expected utility is lower than that obtained in the strategic bank problem. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$, $\alpha = 0.15$ and $\varepsilon = 0.1$.



patible (and thus where the allocation is different under asymmetric information), is also for low probability values, but ones that are large enough for the no-default regime to be superior (see appendix A.4). The numerical results show that this region of intermediate probability is very small.

Figure 3.6 abstracts from default to highlight the differences between the full and asymmetric information allocations of the the global bank, and shows that the differences, when they exist, become smaller as p increases. Thus in the small region where there are differences between the global bank allocations under full and asymmetric information, the differences in allocations are predicted to be small, leading to negligible expected utility effects.

The top-right panel of Figure 6.2 shows that the expected utility in the no-default equilibrium is always weakly lower than the expected utility of the fully informed global bank allocation. The same holds in mutual-default equilibrium in the bottom-right panel of the figure. However, the bottom-left panel of the figure shows that the expected utility in the single-default equilibrium is often larger than the expected utility of the fully informed global bank allocation.

In Table 6.1 I present descriptive statistics on the distribution of computed differences in expected utility with respect to the least constrained benchmark, the global bank under full information. I denote the expected utility of the global bank allocation under full information by EU^{GB^F}. In the table I give descriptive statistic on the quantities $EU^i - EU^{GB^F}$, where the components, EU^i, I consider are the expected utility of the global bank allocation under asymmetric information, EU^{GB^A}, and the equilibrium expected utility in the three types of Nash equilibrium of the strategic bank problem, EU_j^{NE} (where $j = \{nd, sd, md\}$ for the no-default, single-default and mutual-default equilibria respectively). I first consider the full distribution of the difference between each comparison (i.e. the distribution of $EU^i - EU^{GB^F}$) and then split the distribution into its positive and negative parts (i.e. the distributions of $EU^i - EU^{GB^F} > 0$ and $EU^i - EU^{GB^F} < 0$).

The positive part of the distribution shows that in 71.85% of cases ($\frac{1062}{1478}$), the single-default equilibrium yields expected utility larger than that of the global bank allocation under full information. Moreover as shown in Figure 6.3, these positive utility differences are sizeable – frequently more that 5 percentage points. In contrast a very small number of cases for the global bank allocation under asymmetric information ($\frac{6}{10207} = 0.05\%$), the no-default equilibrium ($\frac{36}{8066} = 0.44\%$) and the mutual-default equilibrium ($\frac{17}{663} = 2.56\%$) yield expected utility greater than that of the global bank allocation under full information. In addition, the average sizes of these differences are negligible[3].

The negative part of the distribution in Table 6.1 is more ambiguous, so as a final piece of evidence I present t tests of the hypothesis that the mean of the expected utility in the global bank allocation under full information is equal to each of the other allocations against the al-

[3]I attribute these discrepancies to numerical imprecision

Figure 6.2: Expected utility comparison between the global bank allocation under full information and four other cases (the global bank allocation under asymmetric information and the three different types of Nash equilibria) across computational results for 10 207 unique parameter vectors that cover the full parameter space. In each case, the expected utility of the global bank under full information is plotted on the horizontal axis. On the vertical axes are the expected utility of: the global bank allocation under asymmetric information (top left); the no-default equilibrium (top right); the single-default equilibrium (bottom left); the mutual-default equilibrium (bottom right).

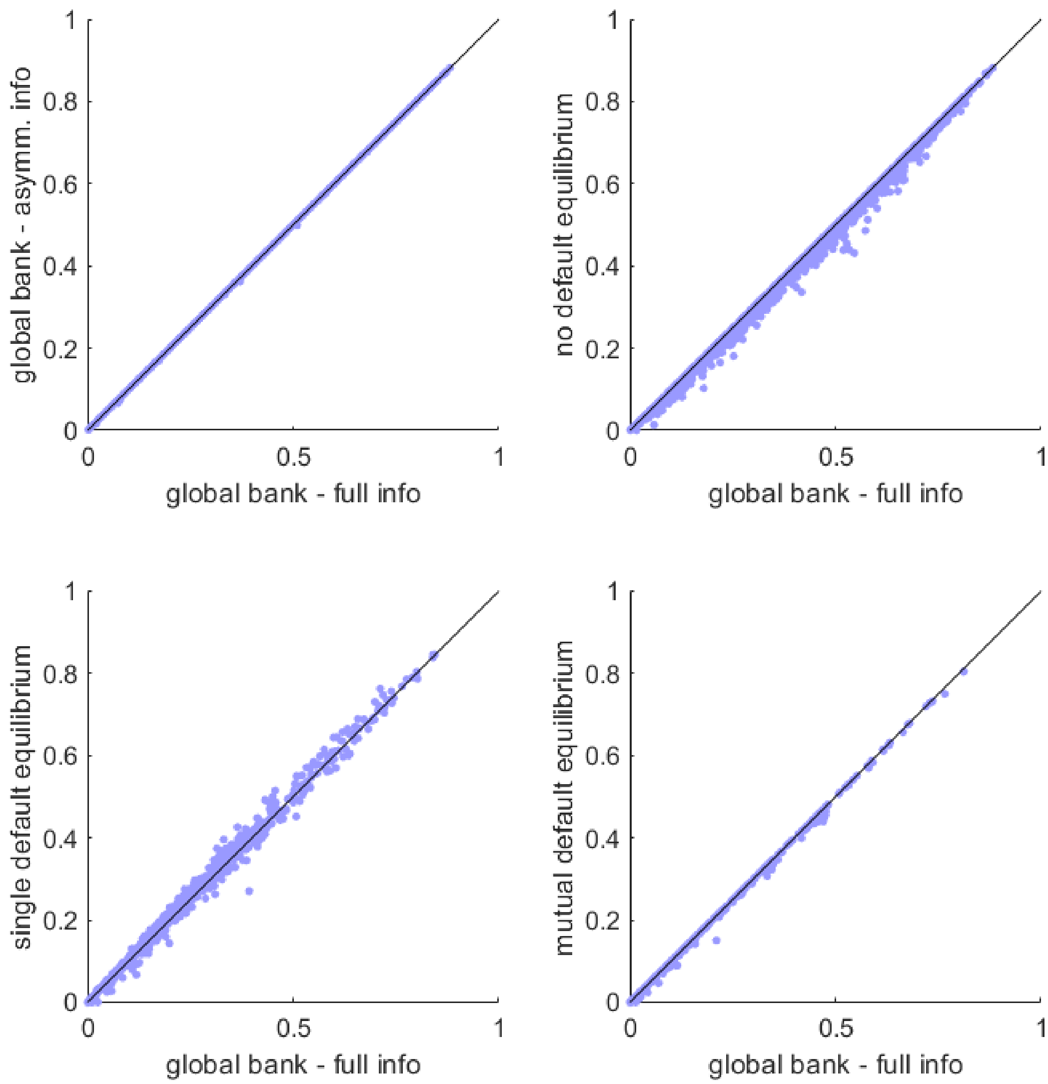
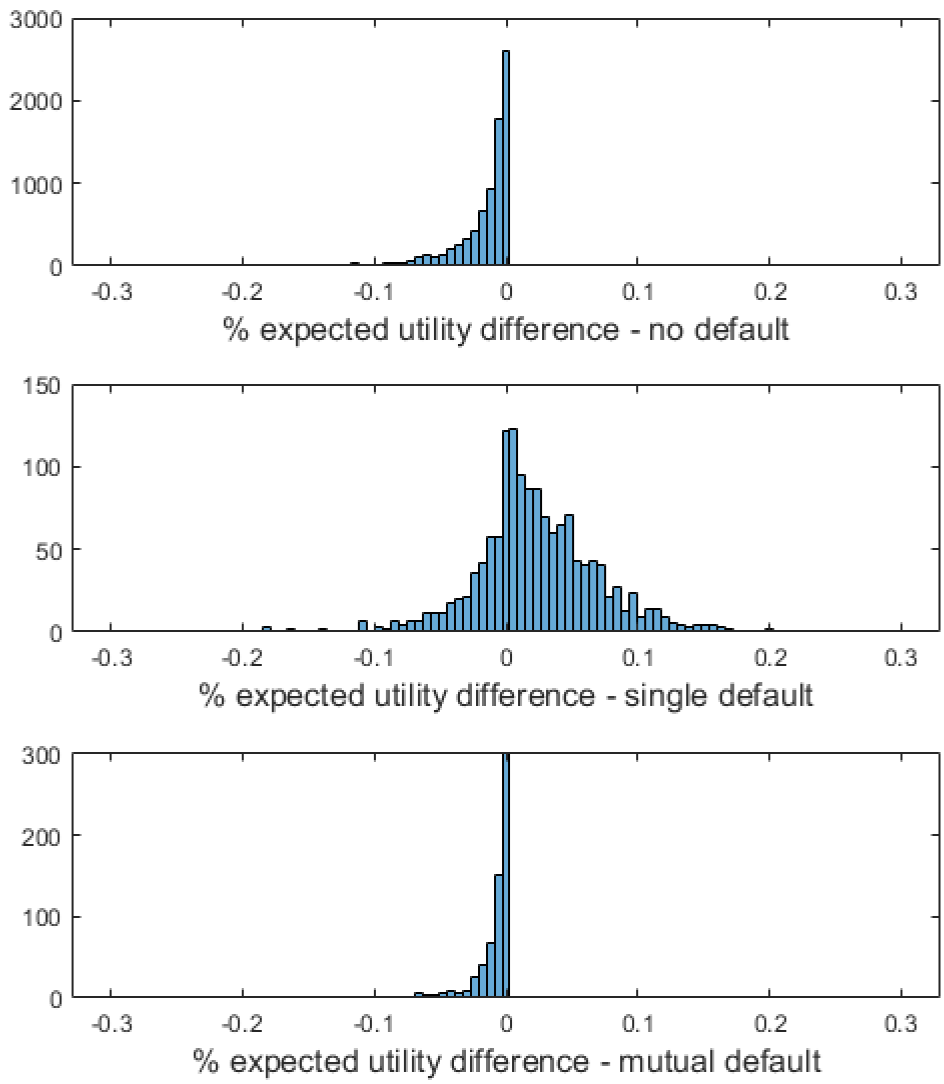


Table 6.1: Descriptive statistics of the distributions of expected utility differences of various allocations with that of the global bank allocation under full information, EU^{GB^F}. I consider the distributions of differences in expected utility with respect to four allocations: the global bank under asymmetric information ($EU^{GB^A} - EU^{GB^F}$), the no-default equilibrium ($EU^{NE}_{nd} - EU^{GB^F}$), the single-default equilibrium ($EU^{NE}_{sd} - EU^{GB^F}$), and the mutual-default equilibrium ($EU^{NE}_{md} - EU^{GB^F}$). The table gives the full distributions (top panel), the positive parts (middle panel) and the negative parts (bottom panel) of the distributions.

	Count	Min	Max	Mean	Median	Std. Dev.	Skewness
Full distribution							
$EU^{GB^A} - EU^{GB^F}$	10207	- 0.0107	0.0002	0	0	0.0002	-30.56
$EU^{NE}_{nd} - EU^{GB^F}$	8066	- 0.1148	0.0043	-0.0042	-0.0015	0.0073	-4.061
$EU^{NE}_{sd} - EU^{GB^F}$	1478	- 0.1226	0.0653	0.0043	0.0029	0.0123	-0.556
$EU^{NE}_{md} - EU^{GB^F}$	663	- 0.0594	0.0003	-0.0026	-0.0008	0.0048	-4.615
Positive part							
$EU^{GB^A} - EU^{GB^F} > 0$	6	0	0.0002	0	0	0	1.784
$EU^{NE}_{nd} - EU^{GB^F} > 0$	36	0	0.0043	0.0002	0	0.0007	5.334
$EU^{NE}_{sd} - EU^{GB^F} > 0$	1062	0	0.0653	0.009	0.0063	0.0092	2.134
$EU^{NE}_{md} - EU^{GB^F} > 0$	17	0	0.0003	0	0	0.0001	1.851
Negative part							
$EU^{GB^A} - EU^{GB^F} \leq 0$	10201	- 0.0107	0	0	0	0	-30.56
$EU^{NE}_{nd} - EU^{GB^F} \leq 0$	8030	- 0.1148	0	-0.0042	-0.0015	0.0073	-4.055
$EU^{NE}_{sd} - EU^{GB^F} \leq 0$	416	- 0.1226	0	-0.0077	-0.0037	0.0109	-4.109
$EU^{NE}_{md} - EU^{GB^F} \leq 0$	646	- 0.0594	0	-0.0027	-0.0009	0.0048	-4.573

Figure 6.3: Histograms of expected utility differences of the no-default, single-default and mutual-default equilibria with that of the global bank allocation under full information, presented as percentages of the expected utility of the global bank allocation under full information.



ternative that it is not equal. The tests that the mean of the expected utility of the global bank allocation under full information is equal to that of the other allocations cannot be rejected (at the 5% level) for the global bank under asymmetric information (p value 0.9965), the no-default equilibrium (p value 0.3860) or the mutual-default equilibrium (p value 0.6056). However, the hypothesis is rejected for the single-default equilibrium (p value 0.0327).

I conclude by summarizing the findings in this section as my final numerical result:

Numerical Result 9. *The numerical comparison between the expected utility of the aggregate benchmarks and the Nash equilibrium of the non-cooperative game between two banks yields the following conclusions:*

(i) Whenever the Nash equilibrium of the game is characterized by no default or mutual default, there is no evidence that the decentralized outcome is welfare inferior to the benchmarks.

(ii) When the Nash equilibrium of the game is characterized by single default, however, there is evidence that the decentralized equilibrium can be superior to the aggregate benchmarks.

Conclusion

This dissertation presents a model of the mechanisms of contagion among financial intermediaries. Several novel results are contributed to the literature, specifically to the strand focusing on contagion among deposit taking banks following Diamond and Dybvig (1983), Allen and Gale (2000), Freixas et al. (2000) and Castiglionesi (2007).

The economy in this dissertation is subject to pure liquidity risk, in the form of both regional and aggregate liquidity demand shocks. Liquidity risk often plays a central role in banking crises, including the global financial crisis of 2007/8 (Brunnermeier, 2009; Gorton and Metrick, 2012; Copeland et al., 2014; Krishnamurthy et al., 2014; Martin et al., 2014). Many papers have considered contagion in the face of aggregate liquidity risk combined with other types of risk, but in the context of pure liquidity risk, this work is the first to both study the sources of contagion and quantify the prevalence of contagion in the full parameter space of a model of strategic banks.

The model economy has two regions, each with a local bank. There are three periods and two types of assets: liquidity and a productive investment. Liquidity is a storage technology that transfers the consumption good from one period to the next at zero net return. The productive asset has a higher return than liquidity if held to maturity (until the final period), but a lower return than liquidity if liquidated early (in the intermediate period). All investment decisions occur in the first period, before any shock realizes. In the intermediate period, shocks realize and all uncertainty is resolved, except that individual liquidity preferences remain private information.

Consumers in each region are subject to liquidity demand uncertainty. Consumers do not know whether they will need liquidity to consume in the short run (in the intermediate period, labelled *early consumers*) or whether they could use the productive investment to consume more in the long run (in the final period, labelled *late consumers*). This liquidity demand uncertainty creates a role for a bank, which can offer insurance to consumers against their liquidity demand risk by pooling resources of all regional consumers and offering a consumption allocation that is superior to the allocation that consumers without access to a bank could obtain.

Banks are defined as financial intermediaries that offer demand deposit contracts. The demand deposit contract of a bank provides a risk-free return to early consumers, except when the bank defaults due to a bank run (which occurs when all depositors, early and late, attempt to withdraw their deposits in the intermediate period). If a bank does default, it is fully liquidated and all proceeds are distributed pro rata to all claimants. Banks in this literature are typically also assumed to be subject to an information constraint: they cannot observe the type of consumer (early or late) that withdraws in the intermediate period.

From the perspective of a regional bank there are two types of liquidity demand shocks, which induce randomness in the proportion of early and late consumers. First, there is a regional liquidity demand shock that is perfectly negatively correlated across regions, which does not impact the average global liquidity demand. The regional liquidity demand shock provides an incentive for a regional bank to insure itself against the regional liquidity risk by holding a positive amount of the deposit contract of the bank in the other region. Second, there is an aggregate liquidity demand shock which increases the average global liquidity demand. This shock may hit either region with equal probability. The aggregate liquidity demand shock may cause contagion, which is defined as follows: contagion occurs when the bank hit by the aggregate liquidity demand shock defaults, and the liquidation value of that bank is so low that the counter party bank also fails. This is due to the reduction in the value of the claims on the interbank deposit of the bank in the region hit by the shock. My contribution to the literature lies in characterizing the consequences of an aggregate liquidity demand shock across the full range of its probability, when that probability is common knowledge (Allen and Gale, 2000;

Castiglionesi, 2007).

In this model economy, I present a novel aggregate benchmark allocation that provides a welfare comparison for the outcomes that are obtained in the decentralized equilibrium. This benchmark is labelled a global bank and is motivated as a single bank that operates a branch in each of the two regions. This means that consumers that deposit at the regional branches are treated identically. The global benchmark is studied in two cases. The first case is a global bank with full information, where the global bank can observe the type of consumer that withdraws in the intermediate period (i.e. the bank is not subject to the typical information constraint imposed on banks). In this case, the bank can prevent a run by assumption. The second case is a global bank under asymmetric, where the bank cannot observe the type of consumer that withdraws in the intermediate period. In this case, the global bank must offer an incentive compatible consumption allocation to avoid a run by late consumers in the intermediate period.

The global bank is always subject only to the constraint of having to offer an early consumption level that is non-state-contingent except if the bank defaults. Default in the intermediate period is therefore benchmark equivalent of contagion in the decentralized case. It is the only tool in this model whereby a global bank can transfer consumption risk from late consumers to early consumers.

The benchmark allocation is novel in the sense that it combines two essential features that have not been simultaneously imposed on aggregate benchmarks in the literature. These features are that (i) the global bank benchmark is constrained to offer non-state-contingent early consumption in the absence of default (which is novel relative to e.g. Castiglionesi et al. (2017)), and that (ii) the benchmark allows default as a potentially optimal outcome (this is novel relative to, for example Allen et al. (2009)).

The first benchmark contribution is to show that, even in a model with only liquidity demand risk, default (and therefore contagion) can be an optimal outcome. This is similar to the result in Allen and Gale (1998), except that they considered a setting with competitive banks that face both liquidity and investment risk. The only constraint necessary for this result is that early consumption must be non-state-contingent in the absence of default. This adds to the results of Allen et al. (2009) and Castiglionesi et al. (2017) who do not allow contagion to be potentially optimal.

The second benchmark contribution is to show that, in the absence of default, a global bank achieves an optimal consumption allocation by balancing the ex post inefficiencies of two tools: excess liquidity or partial liquidation of the investment. Excess liquidity occurs when the the aggregate liquidity demand turns out to be small, so that late consumption is partially funded out of liquidity rather than solely out of the productive investment, which would have been more efficient ex post. Partial liquidation of the productive investment occurs when the aggregate liquidity demand turns out to be large, so that early consumption is funded partially by the ex post inefficient early liquidation of the productive investment.

The final benchmark result shows that the information constraint – which distinguishes the global bank allocations under full and asymmetric information – is important. The differences between the benchmark allocations are characterized, as well as the parameter regions where they differ. The results show that, when the allocations differ, a global bank under asymmetric information can offer less liquidity risk insurance to consumers than under full information.

The most important results relate to the decentralized problem. These results characterize the outcomes when two non-cooperative banks choose their portfolios and deposit returns independently. The solution concept is the Nash equilibrium of the strategic game between two representative banks, one in each of the two regions.

The decentralized problem inherits the constraints from the benchmarks: deposit returns must be risk free (except in the case of bank default), and the contracts offered must be incentive compatible in order to avoid a bank run (i.e. regional banks cannot observe the consumer types). In the aggregate benchmark, the global bank aggregates across regions, which reduces the problem to one with only two states and two potentially binding incentive compatibility constraints due to the information asymmetry. In this situation, the benchmark problems are amenable to analytic solutions. In the decentralized problem, however, there are four states and several potentially binding constraints. This necessitates a numerical approach, as a fully analytical approach is infeasible.

The results for the symmetric decentralized banking equilibrium are computational, given that a fully analytic treatment is infeasible. The approach is to use a numeric algorithm to solve for the Nash equilibrium of the game between two strategic banks facing both regional and aggregate liquidity demand shocks. The algorithm finds the numerically approximated fixed point of the best response choice of portfolio and deposit return of the bank in one region in response to the choice of the bank in another region.

The computational approach yields several results that are novel in the literature on financial contagion. In a broader sense, one of the general contributions of the dissertation is to show that the numerical/computational study of theoretical problems in the microeconomics of financial intermediation is a productive option. Numerical results are common in macroeconomics, but far less so in this microeconomic literature. This dissertation serves as an example of the power of using computational methods to provide insight into economic situations that are not amenable to purely analytic methods.

The numerical results on the decentralized banking problem can be summarized as follows. First, there is no evidence of asymmetric equilibria in this game, even though the algorithm is robust to finding them. Second, as the probability of aggregate risk approaches zero, the symmetric Nash equilibrium converges on the deterministic first-best allocation of Allen and Gale (2000), within numerical precision. Third, the symmetric Nash equilibrium can have three different characterizations: if the probability of the aggregate liquidity demand shock is small enough and the shock is large enough, the equilibrium is characterized by both banks defaulting if the aggregate shock realizes in either region (i.e. contagion occurs). If the probability is intermediate, the equilibrium is characterized by only one bank defaulting (the one hit by the aggregate liquidity demand shock). Thus, a bank default occurs, but without causing contagion. If the probability of the aggregate liquidity demand shock is high enough, the banks choose symmetric portfolios and deposit returns such that neither bank ever defaults. Finally, at the parameter boundary where one equilibrium type switches to another, there are discontinuities in the optimal choices. This mirrors similar discontinuities in the aggregate benchmark allocations when the solution switches from one regime to another (i.e. when the optimal choice switches from default when the aggregate liquidity demand shock hits, to no default).

Arguably the most important result is related to the overall prevalence of financial contagion. When banks are able to internalize the aggregate liquidity demand risk, financial contagion is a very rare phenomenon in this model. In an exercise where parameter sets were drawn at random from their possible ranges, contagion was only *possible* in approximately 4% of the draws. Moreover, one of these randomly drawn parameters is the probability of an aggregate liquidity demand shock, which must realize for contagion to occur. Taking into account the probability of an aggregate liquidity demand shock occurring, the ex-ante likelihood of contagion falls to approximately 0.5%.

The final results in this dissertation are a set of welfare comparisons between the Nash equilibrium of the decentralized banking problem and the solutions to the aggregate benchmarks. When the Nash equilibrium is characterized either by no default or by mutual default (contagion), there is no evidence that the decentralized equilibrium is inferior to the aggregate benchmarks. The differences that are there may be due to numerical imprecision. However, when the Nash equilibrium is characterized by single default (i.e. default without contagion), there is strong evidence that the aggregate benchmarks are inferior to the solution in the decentralized case. This is tangentially similar to the results on the inefficient integration or diversification of banks found by Wagner (2010) and Castiglionesi et al. (2017) in different model settings. The fact that the global bank aggregates over regions and treats consumers in different regions identically, means that default can only be applied in both regions or neither region. In the decentralized case, there is an additional degree of freedom as it is possible for only one bank to default while the other survives.

The fact that the aggregate benchmarks are never superior to the decentralized outcomes, and sometimes inferior, means there is little that the model can say about simple, central-planning type policies to improve outcomes in this model. Indeed, since contagion is optimal when it occurs, any aggregate policies, such as forcing banks to hold more liquidity to avoid contagion, will be welfare reducing. The next direction that research could take, therefore, is to consider whether other policies (such as deposit insurance) can improve on welfare, but this will require a different benchmark as motivation for the possibility of welfare improvement. For instance, the central assumption of non-contingent deposit returns will probably have to be relaxed as in Castiglionesi (2007) and Castiglionesi et al. (2017).

Other extensions are also likely to yield important results. First, in this dissertation the two regions are fully symmetric. If the regions were asymmetric, it would be possible to study the effects of the interactions between smaller and larger regional banks. For instance, one could ask whether a smaller bank could cause contagious effects on a larger bank. Second, the model only considered two regions in order to keep the analysis of the decentralized game tractable, but since the solutions were obtained computationally, it should be possible to extend the solution algorithm to include more banks. This would allow for a richer study of endogenous interbank network formation. Finally, for tractability and for comparability to the seminal model of Allen and Gale (2000), the deposit contract treated banks and consumers identically. A more realistic extension would be to make the interbank deposit market distinct from the consumer deposit contract, in both return and seniority. The study of the simultaneous solution to the banking problem with a fully strategic interbank market structure would be a valuable approach to modelling features like over-the-counter interbank market transactions. I will consider these avenues in future work.

Bibliography

Acemoglu, D., Ozdaglar, A. and Tahbaz-Salehi, A. (2015), 'Systemic risk and stability in financial networks', *American Economic Review* **105**(2), 564–608.

Acharya, V. V., Shin, H. S. and Yorulmazer, T. (2009), 'Endogenous choice of bank liquidity: the role of fire sales', *Bank of England Working Paper, WP376* .

Acharya, V. V. and Skeie, D. (2011), 'A model of liquidity hoarding and term premia in inter-bank markets', *Journal of Monetary Economics* **58**(5), 436–447.

Acharya, V. V. and Yorulmazer, T. (2008), 'Cash-in-the-market pricing and optimal resolution of bank failures', *Review of Financial Studies* **21**(6), 2705–2742.

Afonso, G., Kovner, A. and Schoar, A. (2011), 'Stressed, not frozen: The federal funds market in the financial crisis', *The Journal of Finance* **66**(4), 1109–1139.

Ahnert, T. and Georg, C.-P. (2018), 'Information contagion and systemic risk', *Journal of Financial Stability* **35**, 159–171.

Allen, F. and Babus, A. (2009), 'Networks in finance', *in* P. R. Kleindorfer and Y. J. Winds, eds, '*The network challenge: strategy, profit, and risk in an interlinked world*', Wharton School Publishing Upper Saddle River, NJ, pp. 367–382.

Allen, F., Babus, A. and Carletti, E. (2012), 'Asset commonality, debt maturity and systemic risk', *Journal of Financial Economics* **104**, 519–534.

Allen, F., Carletti, E. and Gale, D. (2009), 'Interbank market liquidity and central bank intervention', *Journal of Monetary Economics* **56**(5), 639–652.

Allen, F. and Gale, D. (1998), 'Optimal financial crises', *Journal of Finance* **53**(4), 1245–1284.

Allen, F. and Gale, D. (2000), 'Financial contagion', *Journal of Political Economy* **108**, 1–33.

Allen, F. and Gale, D. (2004*a*), 'Financial fragility, liquidity, and asset prices', *Journal of the European Economic Association* **2**(6), 1015–1048.

Allen, F. and Gale, D. (2004*b*), 'Financial intermediaries and markets', *Econometrica* **72**(4), 1023–1061.

Arciero, L., Heijmans, R., Heuver, R., Massarenti, M., Picillo, C. and Vacirca, F. (2016), 'How to Measure the Unsecured Money Market? The Eurosystem's Implementation and Validation Using TARGET2 Data', *International Journal of Central Banking* **12**, 247–280.

Babus, A. (2016), 'The formation of financial networks', *The RAND Journal of Economics* **47**(2), 239–272.

Bala, V. and Goyal, S. (2000), 'A noncooperative model of network formation', *Econometrica* **68**(5), 1181–1229.

Bloch, F. and Jackson, M. O. (2006), 'Definitions of equilibrium in network formation games', *International Journal of Game Theory* **34**(3), 305–318.

Boser, B. E., Guyon, I. M. and Vapnik, V. N. (1992), 'A training algorithm for optimal margin classifiers', *in* '*Proceedings of the fifth annual workshop on Computational learning theory*', ACM, pp. 144–152.

Brunnermeier, M. K. (2009), 'Deciphering the liquidity and credit crunch 2007-2008', *Journal of Economic perspectives* **23**(1), 77–100.

Brunnermeier, M. K. and Pedersen, L. H. (2008), 'Market liquidity and funding liquidity', *The review of financial studies* **22**(6), 2201–2238.

Caballero, R. J. and Krishnamurthy, A. (2008), 'Collective risk management in a flight to quality episode', *The Journal of Finance* **63**(5), 2195–2230.

Cabrales, A., Gale, D. and Gottardi, P. (2015), 'Financial contagion in networks', *European University Institute Working Paper ECO2015/01* .

Calomiris, C. W. and Carlson, M. (2016), 'Corporate governance and risk management at unprotected banks: National banks in the 1890s', *Journal of Financial Economics* **119**(3), 512–532.

Calomiris, C. W. and Carlson, M. (2017), 'Interbank networks in the national banking era: their purpose and their role in the panic of 1893', *Journal of Financial Economics* **125**(3), 434–453.

Castiglionesi, F. (2007), 'Financial contagion and the role of the central bank', *Journal of Banking & Finance* **31**(1), 81–101.

Castiglionesi, F., Feriozzi, F. and Lorenzoni, G. (2017), 'Financial Integration and Liquidity Crisis', *Management Science* **65**(3), 955–975.

Copeland, A., Martin, A. and Walker, M. (2014), 'Repo runs: Evidence from the tri-party repo market', *The Journal of Finance* **69**(6), 2343–2380.

Cortes, C. and Vapnik, V. (1995), 'Support-vector networks', *Machine learning* **20**(3), 273–297.

Dasgupta, A. (2004), 'Financial contagion through capital connections: A model of the origin and spread of bank panics', *Journal of the European Economic Association* **2**(6), 1049–1084.

de Bandt, O. and Hartmann, P. (2000), 'Systemic risk: a survey', *European Central Bank Working Paper, No. 35* .

Diamond, D. W. and Dybvig, P. H. (1983), 'Bank runs, deposit insurance and liquidity', *Journal of Political Economy* **91**, 401–419.

Diamond, D. W. and Rajan, R. G. (2001), 'Liquidity risk, liquidity creation, and financial fragility: A theory of banking', *Journal of political Economy* **109**(2), 287–327.

Diamond, D. W. and Rajan, R. G. (2005), 'Liquidity shortages and banking crises', *The Journal of finance* **60**(2), 615–647.

Diamond, D. W. and Rajan, R. G. (2011), 'Fear of fire sales, illiquidity seeking, and credit freezes', *Quarterly Journal of Economics* **126**(2), 557–591.

Duarte, F. and Eisenbach, T. M. (2018), 'Fire-sale spillovers and systemic risk', *Federal Reserve Board of New York, Staff Report* (645).

Dutta, B. and Jackson, M. O. (2003), 'On the formation of networks and groups', *in* B. Dutta and M. O. Jackson, eds, '*Networks and groups*', Springer, Berlin, Heidelberg, pp. 1–15.

Eisenberg, L. and Noe, T. H. (2001), 'Systemic risk in financial systems', *Management Science* **47**(2), 236–249.

Eisfeldt, A. and Rampini, A. (2006), 'Capital reallocation and liquidity', *Journal of Monetary Economics* **53**(3), 369–399.

Elliott, M., Golub, B. and Jackson, M. O. (2014), 'Financial networks and contagion', *American Economic Review* **104**(10), 3115–3153.

Ennis, H. M. and Keister, T. (2006), 'Bank runs and investment decisions revisited', *Journal of Monetary Economics* **53**(2), 217–232.

Flannery, M. J. (1996), 'Financial crises, payment system problems, and discount window lending', *Journal of money, credit and banking* **28**(4), 804–824.

Forbes, K. (2012), 'The "big c": identifying contagion', *National Bureau of Economic Research Working Paper w18465* .

Freixas, X. and Jorge, J. (2008), 'The role of interbank markets in monetary policy: A model with rationing', *Journal of Money, Credit and Banking* **40**(6), 1151–1176.

Freixas, X., Martin, A. and Skeie, D. (2011), 'Bank liquidity, interbank markets, and monetary policy', *The Review of Financial Studies* **24**(8), 2656–2692.

Freixas, X., Parigi, B. M. and Rochet, J. C. (2000), 'Systemic risk, interbank relations, and liquidity provision by the central bank', *Journal of Money, Credit and Banking* **32**(3), 611–638.

Friedman, M. and Schwartz, A. J. (2008), *A monetary history of the United States, 1867-1960*, Princeton University Press, Princeton, NJ.

Fung, G. M. and Mangasarian, O. L. (2005), 'Multicategory proximal support vector machine classifiers', *Machine learning* **59**(1-2), 77–97.

Furfine, C. H. (2003), 'Interbank exposures: Quantifying the risk of contagion', *Journal of Money, Credit, and Banking* **35**, 111–128.

Gai, P., Haldane, A. and Kapadia, S. (2011), 'Complexity, concentration and contagion', *Journal of Monetary Economics* **58**(4), 453–470.

Gai, P. and Kapadia, S. (2010), 'Contagion in financial networks', *Proceedings of the Royal Society A* **466**, 2401–2423.

Glasserman, P. and Young, H. P. (2015), 'How likely is contagion in financial networks?', *Journal of Banking & Finance* **50**, 383–399.

Goldstein, I. and Pauzner, A. (2005), 'Demand deposit contracts and the probability of bank runs', *Journal of Finance* **60**(3), 1293–1327.

Gorton, G. and Metrick, A. (2012), 'Securitized banking and the run on repo', *Journal of Financial economics* **104**(3), 425–451.

Gromb, D. and Vayanos, D. (2010), 'Limits of arbitrage', *Annual Review of Financial Economics* **2**(1), 251–275.

Heider, F., Hoerova, M. and Holthausen, C. (2015), 'Liquidity hoarding and interbank market rates: The role of counterparty risk', *Journal of Financial Economics* **118**(2), 336–354.

Jackson, M. O. and Wolinsky, A. (1996), 'A strategic model of social and economic networks', *Journal of economic theory* **71**(1), 44–74.

Krishnamurthy, A., Nagel, S. and Orlov, D. (2014), 'Sizing up repo', *The Journal of Finance* **69**(6), 2381–2417.

Laux, C. and Leuz, C. (2010), 'Did fair-value accounting contribute to the financial crisis?', *Journal of economic perspectives* **24**(1), 93–118.

Martin, A., Skeie, D. and Thadden, E.-L. v. (2014), ‘Repo Runs’, *The Review of Financial Studies* **27**(4), 957–989.

Matlab (2017), ‘Matlab, version r2017a’, *MathWorks, Natick, MA, USA* **https://www.mathworks.com/**.

Milgrom, P. and Segal, I. (2002), ‘Envelope theorems for arbitrary choice sets’, *Econometrica* **70**(2), 583–601.

Minoiu, C. and Reyes, J. A. (2013), ‘A network analysis of global banking: 1978–2010’, *Journal of Financial Stability* **9**(2), 168–184.

Mitchener, K. J. and Richardson, G. (2013), ‘Shadowy banks and financial contagion during the Great Depression: a retrospective on Friedman and Schwartz’, *American Economic Review* **103**(3), 73–78.

Mitchener, K. J. and Richardson, G. (2019), ‘Network contagion and interbank amplification during the great depression’, *Journal of Political Economy* **127**(2), 465–507.

Morris, S. and Shin, H. S. (2003), ‘Global Games: Theory and Applications’, *in* M. Dewatripont, L. P. Hansen and S. J. Turnovsky, eds, ‘*Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress*’, Vol. 1 of *Econometric Society Monographs*, Cambridge University Press, Cambridge, UK, p. 56–114.

Rogers, L. C. and Veraart, L. A. (2013), ‘Failure and rescue in an interbank network’, *Management Science* **59**(4), 882–898.

Roukny, T., Battiston, S. and Stiglitz, J. E. (2018), ‘Interconnectedness as a source of uncertainty in systemic risk’, *Journal of Financial Stability* **35**, 93–106.

United States Financial Crisis Inquiry Commission (2011), *Final Crisis Inquiry Report*, United States Government.

Wagner, W. (2010), ‘Diversification at financial institutions and systemic crises’, *Journal of Financial Intermediation* **19**(3), 373–386.

Appendix A

Proofs

I characterize several benchmarks in terms of the bounds on the probability of the aggregate liquidity shock (p) summarized in Table A.1. These bounds separate regions with distinct solution characterizations.

Table A.1: Summary of bounds on p (probability of aggregate liquidity shock)

Notation	Benchmark Setting and Regions Separated
$\underline{p}_{ND}^{GB^F}$	If $p < \underline{p}_{ND}^{GB^F}$, the global bank with full information (GB^F) chooses zero excess liquidity in the no-default regime (ND). Otherwise, excess liquidity is positive.
$\overline{p}_{ND}^{GB^F}$	If $p > \overline{p}_{ND}^{GB^F}$, the global bank with full information chooses zero partial liquidation in the no-default regime. Otherwise, partial liquidation is positive.
$\underline{p}_{D}^{GB^F}$	If $p < \underline{p}_{D}^{GB^F}$, the global bank with full information chooses zero excess liquidity in the default regime (D). Otherwise, excess liquidity is positive.
$\overline{p}_{D}^{GB^F}$	If $p > \overline{p}_{D}^{GB^F}$, the global bank with full information chooses zero investment in the default regime. Otherwise, investment is positive.
$\breve{p}$	If $p < \breve{p}$, the global bank with full information chooses to default in state H, otherwise it chooses not to default in state H.
$\hat{p}_{IC}$	If $\breve{p} < p < \hat{p}_{IC}$, the allocation of a global bank with full information is not incentive compatible (IC) and, hence, the allocation of a global bank with asymmetric information differs from that of a fully informed global bank. Otherwise, the two allocations are identical.
$\underline{p}_{ND}^{GB^A}$	If $\breve{p} < p < \hat{p}_{IC}$ and $p < \underline{p}_{ND}^{GB^A}$, a global bank under asymmetric information (GB^A) chooses zero excess liquidity. Otherwise, if $\breve{p} < p < \hat{p}_{IC}$ and $p > \underline{p}_{ND}^{GB^A}$, a global bank under asymmetric information chooses positive excess liquidity and partial liquidation.
$\underline{p}^{Aut}, \overline{p}^{Aut}$	If $p' = \gamma + p\alpha < \underline{p}^{Aut}$, a consumer in autarky (Aut) invests fully. If $p' > \overline{p}^{Aut}$, a consumer in autarky does not invest at all. Otherwise, if $\underline{p}^{Aut} < p < \overline{p}^{Aut}$, the allocation is interior.

A.1 Proof of Proposition 1

The problem of a global bank that can observe consumer types who is constrained from default in state H, is to choose a consumption allocation to maximize ex ante expected utility of an arbitrary region subject to a non-state-contingency constraint on early consumption: $c_{1L} = c_{1H} \equiv c_1$. The consumption allocation need not be incentive compatible, as the global bank can observe consumer types and can disallow a run by assumption.

In principle, a fully informed global bank that does not default in state S can use any combination of liquidity or investment to fund any part of the necessary consumption allocation, c_1, c_{2L} and c_{2H}. However, it was shown in the text that it is never optimal for such a bank to use partial liquidation of investment to fund c_1 in state L, nor to use liquidity to fund any part of c_{2H}. Thus, it may be optimal to use excess liquidity e to partially fund c_{2L}, or to use partial liquidation of investment λ to partially fund c_1 in state H, or both.

Using the accounting identities, $e \equiv y - \gamma c_1$ and $\lambda \equiv \frac{(\gamma+\alpha)c_1 - y}{r}$, I solve for the portfolio choice $y = \frac{\gamma+\alpha}{\alpha}e + \frac{\gamma r}{\alpha}\lambda$ and express the consumption levels (c_1, c_{2L}, c_{2H}) as (linear) functions of (e, λ):

$$c_1(e,\lambda) = \frac{e + r\lambda}{\alpha} \tag{A.1}$$

$$c_{2L}(e,\lambda) = \frac{e + \left(1-y\right)R}{1-\gamma} = \frac{R}{1-\gamma} - \frac{\left(R\left(\frac{\gamma+\alpha}{\alpha}\right) - 1\right)}{1-\gamma}e - \frac{rR\gamma}{\alpha\left(1-\gamma\right)}\lambda \tag{A.2}$$

$$c_{2H}(e,\lambda) = \frac{\left(1-y-\lambda\right)R}{1-\gamma-\alpha} = \frac{R}{1-\gamma-\alpha} - \frac{\left(\frac{\gamma+\alpha}{\alpha}\right)R}{1-\gamma-\alpha}e - \frac{\left(\frac{r\gamma+\alpha}{\alpha}\right)R}{1-\gamma-\alpha}\lambda. \tag{A.3}$$

As a result, I can also express the expected utility in terms of (e, λ):

$$W_{ND}^{GB^F}(e,\lambda) = (\gamma + p\alpha)u(c_1(e,\lambda)) + (1-p)(1-\gamma)u(c_{2L}(e,\lambda)) + p(1-\gamma-\alpha)u(c_{2H}(e,\lambda). \tag{A.4}$$

Inserting the expressions for consumption levels yields an unconstrained problem in choice variables e and $,\lambda$ with Kuhn-Tucker first-order conditions: $\frac{dW_{ND}^{GB^F}}{de}(e^*,\lambda^*) \leqslant 0$ as $e^* \geqslant 0$ and $\frac{\partial W_{ND}^{GB^F}}{\partial\lambda}(e^*,\lambda^*) \leqslant 0$ as $\lambda^* \geqslant 0$:

$$\frac{dW_{ND}^{GB^F}}{de} = (\gamma + p\alpha)u'(c_1) - (1-p)\big(R(\gamma+\alpha)-\alpha\big)u'(c_{2L}) - p(\gamma+\alpha)Ru'(c_{2H}) \tag{A.5}$$

$$\frac{dW_{ND}^{GB^F}}{d\lambda} = (\gamma + p\alpha)u'(c_1) - (1-p)R\gamma u'(c_{2L}) - p\frac{\alpha + r\gamma}{r}Ru'(c_{2H}). \tag{A.6}$$

Optimum at p=0

At $p = 0$, the first-order conditions (A.5) and (A.6) become:

$$\gamma u'(c_1) \leq \big(\gamma R + (R-1)\alpha\big)u'(c_{2L}) \tag{A.7}$$

$$\gamma u'(c_1) \leq \gamma R u'(c_{2L}). \tag{A.8}$$

Since $(R-1)\alpha > 0$, condition (A.7) is slack when (A.8) binds. Thus $e^* = 0$ and $\frac{\partial e^*}{\partial p}(p=0) = 0$. The optimum is characterized by a binding condition (A.8) evaluated at $e^* = 0$ and λ^*:

$$u'\left(\frac{r\lambda^*}{\alpha}\right) = Ru'\left(\frac{R}{1-\gamma}\left[1 - \frac{r\gamma}{\alpha}\lambda^*\right]\right). \tag{A.9}$$

Because of the Inada conditions, and since the left-hand side increases in λ, while the right-hand side decreases in λ, there exists a unique solution $\lambda^* \in \left(0, \frac{\alpha}{r\gamma}\right)$. The bounds are values of λ consistent with zero consumption in the expression of marginal utilities on both sides of (A.9).

The envelope theorem implies that e^* is continuous in p so $\frac{\partial e^*}{\partial p} = 0$ in a neighbourhood of $p = 0$; hence, there exists a $\underline{p}_{ND}^{GB^F} \in (0,1]$ such that the optimum is characterized by $e^* = 0$ and $\lambda^* > 0$ if $0 \leq p \leq \underline{p}_{ND}^{GB^F}$. This optimum is characterized by the following first-order condition:

$$(\gamma + p\alpha)u'\left(\frac{r\lambda^*}{\alpha}\right) = (1-p)R\gamma u'\left(\frac{R}{1-\gamma}\left[1 - \frac{r\gamma}{\alpha}\lambda^*\right]\right) + pR\left(\gamma + \frac{\alpha}{r}\right)u'\left(\frac{R}{1-\gamma-\alpha}\left[1 - \left(1 + \frac{r\gamma}{\alpha}\right)\lambda^*\right]\right), \tag{A.10}$$

which implies the existence of a unique solution $\lambda^* \in \left(0, \frac{\alpha}{\alpha + r\gamma}\right)$. Note that condition (A.10) contains an additional marginal utility term relative to condition (A.9). The argument of the new term implies a lower upper bound on λ than in condition (A.9).

Optimum at p=1

At $p = 1$, the first-order conditions become:

$$(\gamma+\alpha)u'(c_1) \leq (\gamma+\alpha)Ru'(c_{2H}) \tag{A.11}$$

$$(\gamma+\alpha)u'(c_1) \leq \left(\gamma+\frac{\alpha}{r}\right)Ru'(c_{2H}). \tag{A.12}$$

Since $r < 1$, condition (A.12) is slack when condition (A.11) binds. Thus $\lambda^* = 0$ and $\frac{\partial\lambda^*}{\partial p}(p = 1) = 0$. The optimum is characterized by the binding condition (A.11) evaluated at $\lambda^* = 0$ and e^*:

$$u'\left(\frac{e^*}{\alpha}\right) = Ru'\left(\frac{R}{1-\gamma-\alpha}\left[1-\left(\frac{\gamma+\alpha}{\alpha}\right)e^*\right]\right). \tag{A.13}$$

Because of the Inada conditions, and since the left-hand side increases in e, while the right-hand side decreases in e, there exists a unique solution $e^* \in \left(0, \frac{\alpha}{\gamma+\alpha}\right)$. The bounds are values of e consistent with zero consumption in the marginal utilities on both sides of equation (A.13).

The envelope theorem implies that λ^* is continuous in p with $\frac{\partial\lambda^*}{\partial p} = 0$ in a neighbourhood of $p = 1$; hence, there exists a $\overline{p}_{ND}^{GB^F} \in [0,1)$ such that the optimum is characterized by $\lambda^* = 0$ and $e^* > 0$ if $\overline{p}_{ND}^{GB^F} \leq p \leq 1$. This optimum is characterized by the following first-order condition:

$$\begin{aligned}(\gamma+p\alpha)u'\left(\frac{e^*}{\alpha}\right) &= (1-p)\left[R(\alpha+\gamma)-\alpha\right]u'\left(\frac{1}{1-\gamma}\left[R-\left(R\left(1+\frac{\gamma}{\alpha}\right)-1\right)e^*\right]\right) \\ &\quad +pR(\gamma+\alpha)u'\left(\frac{R}{1-\gamma-\alpha}\left[1-\left(1+\frac{\gamma}{\alpha}\right)e^*\right]\right),\end{aligned} \tag{A.14}$$

which implies the existence of a unique solution $e^* \in \left(0, \frac{\alpha}{\alpha+\gamma}\right)$. Note that condition (A.14) contains an additional marginal utility term relative to condition (A.13). The argument of the new term implies a higher upper bound on e than in condition (A.13); therefore, the previous bounds on e prevail.

Unique bounds on the probability of an aggregate liquidity shock

Next, I establish $\frac{\partial e^*}{\partial p} \geq 0$ and $\frac{\partial \lambda^*}{\partial p} \leq 0$, with strict inequality for positive levels of the choice variables. These monotonicity results imply that $\underline{p}_{ND}^{GB^F}$ and $\overline{p}_{ND}^{GB^F}$ are unique.

Case 1: $e^* > 0$ **and** $\lambda^* = 0$

Total differentiation of the first-order condition in (A.5) with respect to p implies that $\frac{\partial e^*}{\partial p} > 0$ whenever

$$u'(c_{2L}) \quad > \quad \frac{R\gamma}{\left(R\gamma + (R-1)\alpha\right)} u'(c_{2H}),$$

for which $u'(c_{2L}) > u'(c_{2H})$ is sufficient. This sufficient condition is satisfied whenever

$$\begin{aligned}
c_{2L}^* &< c_{2H}^* \\
\frac{R}{1-\gamma} - \frac{\left(R\left(\frac{\gamma+\alpha}{\alpha}\right) - 1\right)}{1-\gamma} e^* &< \frac{R}{1-\gamma-\alpha} - \frac{\left(\frac{\gamma+\alpha}{\alpha}\right) R}{1-\gamma-\alpha} e^* \\
e^* &< \tilde{e} \equiv \frac{\alpha R}{(R-1)(\gamma+\alpha)+1}.
\end{aligned}$$

At $e = \tilde{e}$, I obtain $c_1(\tilde{e}) = c_{2L}(\tilde{e}) = c_{2H}(\tilde{e}) = \tilde{c}$ and $\frac{dW_{ND}^{GB^F}}{de}(e = \tilde{e}, \lambda = 0) = -\frac{(\gamma+\alpha)(R-1)}{\alpha} u'(\tilde{c}) < 0$. Thus $e^* < \tilde{e}$ whenever $\lambda^* = 0$ and $\frac{\partial e^*}{\partial p} > 0$ in case 1.

Case 2: $\lambda^* > 0$ **and** $e^* = 0$

Total differentiation of the first-order condition (A.6) implies that $\frac{\partial \lambda^*}{\partial p} < 0$ whenever

$$u'(c_{2H}) > \frac{r(\gamma+\alpha)}{r\gamma+\alpha} u'(c_{2L}), \tag{A.15}$$

for which $u'\left(c_{2H}^*\right) > u'\left(c_{2L}^*\right)$ is sufficient. This sufficient condition is satisfied whenever

$$\begin{aligned}
c_{2H}^* &< c_{2L}^* \\
\frac{R}{1-\gamma-\alpha} - \frac{\left(\frac{r\gamma+\alpha}{\alpha}\right) R}{1-\gamma-\alpha} \lambda^* &< \frac{R}{1-\gamma} - \frac{rR\gamma}{\alpha\left(1-\gamma\right)} \lambda^* \\
1-\gamma-\alpha+r\gamma\lambda^* &> 0,
\end{aligned}$$

which always holds since $\lambda^* \geq 0$ and $1-\gamma-\alpha > 0$. Thus, $\frac{\partial \lambda^*}{\partial p} < 0$ in case 2.

Case 3: $e^* > 0$ **and** $\lambda^* > 0$

Total differentiation of the first-order conditions in (A.5) and (A.6) with respect to p yields

$$\frac{d\lambda^*}{dp} = \frac{N^\lambda}{D} < 0, \quad \frac{de^*}{dp} = \frac{N^e}{D} > 0,$$

since the common denominator is negative,

$$\begin{aligned}
D_1 &= u''(c_1)(\gamma+\alpha p)(1-p)r^2(R-1)^2(1-\alpha-\gamma)u''(c_{2L}) > 0\\
D_2 &= u''(c_1)(\gamma+\alpha p)(1-\gamma)p(1-r)^2R^2u''(c_{2H}) > 0\\
D_3 &= (1-p)pR^2u''(c_{2H})u''(c_{2L})(\gamma(R-r)+\alpha(R-1))^2 > 0\\
D &= -p(1-r)(\gamma+\alpha p)\,(D_1+D_2+D_3) < 0,
\end{aligned}$$

while the numerator of $\frac{\partial e^*}{\partial p}$ is negative,

$$\begin{aligned}
N_1^e &= -(1-\gamma)p^2(1-r)R(\alpha+\gamma)u''(c_{2H})(\alpha+\gamma r) > 0\\
N_2^e &= -\gamma^2(1-p)^2r^2(R-1)(1-\alpha-\gamma)u''(c_{2L}) > 0\\
N_3^e &= -(1-\gamma)(1-r)r^2(R-1)(1-\alpha-\gamma)u''(c_1)(\gamma+\alpha p)^2 > 0\\
N^e &= -u'(c_{2L})\left(R(\gamma(R-r)+\alpha(R-1))\left(N_1^e+N_2^e\right)+N_3^e\right) < 0,
\end{aligned}$$

and the numerator of $\frac{\partial \lambda^*}{\partial p}$ is positive,

$$\begin{aligned}
N_1^\lambda &= -(1-\gamma)p^2(1-r)R^2(\alpha+\gamma)^2u''(c_{2H}) > 0\\
N_2^\lambda &= -\gamma(1-p)^2r(R-1)(1-\alpha-\gamma)u''(c_{2L})(\alpha(R-1)+\gamma R) > 0\\
N_3^\lambda &= -(1-\gamma)(1-r)r(R-1)(1-\alpha-\gamma)(\gamma+\alpha p)^2u''(c_1) > 0\\
N^\lambda &= u'(c_{2L})\left((\gamma(R-r)+\alpha(R-1))(N_1^\lambda+N_2^\lambda)+N_3^\lambda\right) > 0.
\end{aligned}$$

Given these monotonicity results, I can define the following unique bounds on the probability of the aggregate liquidity shock as:

$$\underline{p}_{ND}^{GB^F} \equiv \max\left\{p|e^* = 0\right\}, \quad \overline{p}_{ND}^{GB^F} \equiv \min\left\{p|\lambda^* = 0\right\}.$$

Finally, I show that $\overline{p}_{ND}^{GB^F} > \underline{p}_{ND}^{GB^F}$. The proof is by contradiction. If $\underline{p}_{ND}^{GB^F} \geq \overline{p}_{ND}^{GB^F}$, then there exists a $p'' \in \left[\overline{p}_{ND}^{GB^F}, \underline{p}_{ND}^{GB^F}\right]$ with corresponding optimal choice of $e^* = 0$ and $\lambda^* = 0$. However, this implies $c_1^* = 0$ which contradicts optimality. Thus, $\overline{p}_{ND}^{GB^F} > \underline{p}_{ND}^{GB^F}$. Then, for $p \in \left(\underline{p}_{ND}^{GB^F}, \overline{p}_{ND}^{GB^F}\right)$, the optimum is characterized by $e^* > 0$ and $\lambda^* > 0$ that jointly solve conditions (A.5) and (A.6) when set equal to zero.

A.2 Proof of Proposition 2

The problem facing a global bank that can observe customer types, that is constrained to default and pro-rata pay-out in state H, but which is not constrained to offer non-state-contingent early consumption, is to choose $y \in [0,1]$ and $e \in [0, y]$ to maximize the ex-ante expected utility of consumers subject only to the constraint that $c_{1H} = c_{2H} = y + r(1-y)$.

Using the resource constraint $x + y = 1$ and the definition of excess liquidity in state L, $e = y - \gamma c_{1L} \geq 0$, I can express all levels of consumption in terms of the total liquidity choice y and excess liquidity e in state L:

$$\begin{aligned}
c_{1L}(y,e) &= \frac{y-e}{\gamma} \\
c_{2L}(y,e) &= \frac{e + R(1-y)}{1-\gamma} \\
c_{1H} = c_{2H} \equiv c_H(y,e) &= y + r(1-y).
\end{aligned}$$

As a result, I can also express the expected utility in terms of y and e:

$$W_D^{GB^F}(y,e) = (1-p)\left[\gamma u(c_{1L}(y,e)) + (1-\gamma)u(c_{2L}(y,e))\right] + pu(c_H(y,e)).$$

As in section A.1, the solution to this problem has potential corner solutions. Note that $y = 0$ implies $e = 0$, and thus $c_{1L} = 0$ and cannot be optimal. This yields three remaining cases that characterize the optimal allocation in the default regime of a global bank that can observe consumer types:

Case 1: liquidity only ($y_D^* = 1, e_D^* > 0$)**:**

Since $\frac{dW_D^{GB^F}}{dy}\big|_{p=1} = (1-r)\,u'\big(y + r(1-y)\big) > 0\ \forall y \in [0,1]$, a non-zero measure interval $[\overline{p}_D^{GB^F}, 1]$ must exist such that $y_D^*(p \geq \overline{p}_D^{GB^F}) = 1$. This implies that $c_{2L}^* = c_{1L}^* = c_H^* = 1$ and thus $e_D^*(p \geq \overline{p}_D^{GB^F}) = 1 - \gamma > 0$ in this case.

Next, suppose that $p < \overline{p}_D^{GB^F}$. The Kuhn-Tucker first-order conditions for optimality are as follows: First, for y, the optimality condition is $\frac{dW_D^{GB^F}}{dy}(y_D^*, e_D^*) = 0$, whereby a unique interior solution $y_D^* \in (0,1)$ exists, since the objective function is continuous in liquidity, strictly concave, and satisfies the Inada conditions. The value $y_D^* \in (0,1)$ is implicitly defined by:

$$(1-p)\,u'\left(\frac{y_D^* - e_D^*}{\gamma}\right) + p\,(1-r)\,u'\big(y_D^* + r(1-y_D^*)\big) = (1-p)\,R u'\left(\frac{e_D^* + R(1-y_D^*)}{1-\gamma}\right) \tag{A.16}$$

Second, for e, the Kuhn-Tucker optimality condition is $\frac{\partial W_D^{GB^F}}{\partial e}(y_D^*, e_D^*) \leqslant 0$ as $e \geqslant 0$. Thus, e_D^* is implicitly defined by:

$$u'\left(\frac{e_D^* + R(1-y_D^*)}{1-\gamma}\right) \leqslant u'\left(\frac{y_D^* - e_D^*}{\gamma}\right) \tag{A.17}$$

Case 2: no excess liquidity ($0 < y_D^* < 1, e_D^* = 0$)

At $p = 0$, condition $\frac{dW_D^{GB^F}}{dy}(y_D^*, e_D^*) = 0$ yields $u'(c_{1L}(y_D^*, e_D^*)) = R\,u'(c_{2L}(y_D^*, e_D^*))$.

Since $R > 1$, this implies $u'(c_{1L}(y_D^*, e_D^*)) > u'(c_{2L}(y_D^*, e_D^*))$ so condition (A.17) is slack. Hence, $e_D^*(p = 0) = 0$ and $\frac{\partial e_D^*}{\partial p}(p = 0) = 0$. Since e_D^* is continuous in p, by the theorem of the maximum, there must exist a non-zero measure neighbourhood $[0, \underline{p}_D^{GB^F}]$ such that $e_D^*(p < \underline{p}_D^{GB^F}) = 0$. Moreover, $0 < y_D^* < 1$, $e_D^* = 0$ and $\gamma < 1$ implies $c_{2L}(y_D^*, 0) > c_{1L}(y_D^*, 0) > c_H(y_D^*, 0)$.

Case 3: interior liquidity and excess liquidity $(y_D^* \in (0,1), e_D^* > 0)$

Whenever $e_D^* > 0$ is optimal, which from the above is whenever $p > \underline{p}_D^{GB^F}$, condition (A.17) holds with equality, which implies:

$$\begin{aligned} c_{1L}(y_D^*, e_D^*) = c_{2L}(y_D^*, e_D^*) = c_L^* &= y_D^* + \left(1 - y_D^*\right) R \\ e_D^* &= y_D^* \left(1 + (R-1)\gamma\right) - R\gamma \end{aligned}$$

Moreover, $0 < y_D^* < 1, e_D^* > 0$ and $c_{2L}(y_D^*, e_D^*) = c_{1L}(y_D^*, e_D^*)$ imply that $c_{2L}(y_D^*, e_D^*) = c_{1L}(y_D^*, e_D^*) = y_D^* + R(1 - y_D^*) > 1 > y_D^* + r(1 - y_D^*) = c_H(y_D^*, e_D^*)$.

Establishing uniqueness of bounds

Total differentiation of condition (A.16) with respect to p where $e_D^* > 0$ yields:

$$\frac{dy_D^*}{dp} = -\frac{(1-r)u'(c_H^*) + (R-1)u'(c_L^*)}{p(1-r)^2 u''(c_H^*) + (1-p)(R-1)^2 u''(c_L^*)} > 0 \quad \Rightarrow \frac{de_D^*}{dp} > 0$$

Thus, $\overline{p}_D^{GB^F} > \underline{p}_D^{GB^F}$ and each is unique.

A.3 Proof of Proposition 3

Let $c_{1,ND}$, $c_{2L,ND}$ and $c_{2H,ND}$ be the optimal consumption allocation that solves problem (P1a), the no-default (ND) regime problem of a global bank under full information (GB^F). Therefore, I can state the no-default regime value function as:

$$V_{ND}^{GB^F} \equiv (1-p)\left[\gamma u(c_{1,ND}) + (1-\gamma)u(c_{2L,ND})\right] + p\left[(\gamma+\alpha)u(c_{1,ND}) + (1-\gamma-\alpha)u(c_{2H,ND})\right].$$

Let $c_{1L,D}$, $c_{2L,D}$ and $c_{H,D}$ be the optimal allocation that solves problem (P1b), the default regime (D) problem of a global bank under full information (GB^F). Therefore, I can state the default regime value function as:

$$V_D^{GB^F} \equiv (1-p)\Big[\gamma u(c_{1L,D}) + (1-\gamma)u(c_{2L,D})\Big] + pu(c_{H,D}).$$

At $p = 0$, the fully informed global bank can do no better than the default regime allocation, since the default regime allocation is less constrained in state L than the no-default regime allocation. Even though the outcome in state H has zero weight, the definition of the no-default regime allocation still requires that early consumption be identical across the two states. This imposes a shadow cost on the no-default regime allocation relative to the default regime allocation. Specifically, that $c_{2L,ND} - c_{1,ND} > c_{2L,D} - c_{1L,D}$ – i.e. the there is less liquidity insurance in state L in the no-default regime than in the default regime. In other words, $V_D^{GB^F}\left(p=0\right) \geq V_{ND}^{GB^F}\left(p=0\right)$.

Similarly, the fully informed global bank will not choose the default regime allocation when $p = 1$, as this will imply certain default, and consumption levels less than unity for all consumers for certain. Thus, $V_D^{GB^F}\left(p=1\right) < V_{ND}^{GB^F}\left(p=1\right)$.

Both $V_{ND}^{GB^F}$ and $V_D^{GB^F}$ are continuous in p by the theorem of the maximum. Therefore these value functions must intersect at some $\breve{p} \in [0,1)$.

Next I establish that $\breve{p}$ is unique, and characterize it as a function of the other parameters of the model.

By the envelope theorem that applies in this situation (Milgrom and Segal, 2002):

$$\begin{aligned}
\frac{\partial V_{ND}^{GB^F}}{\partial p} &= \Big[(\gamma+\alpha)u(c_{1,ND}) + (1-\gamma-\alpha)u(c_{2H,ND})\Big] - \Big[\gamma u(c_{1,ND}) + (1-\gamma)u(c_{2L,ND})\Big], \\
\frac{\partial V_{D}^{GB^F}}{\partial p} &= u(c_{H,D}) - \Big[\gamma u(c_{1L,D}) + (1-\gamma)u(c_{2L,D})\Big].
\end{aligned}$$

First note that both of these partial derivatives are non-positive. In other words, in both the default and no-default regimes, the maximum welfare attainable is non-increasing in the probability of the aggregate liquidity demand shock.

In the default regime this is obvious since $c_{H,D} < 1 < c_{1L,D} < c_{2L,D}$. In no-default regime this is less obvious, but still intuitive: when the aggregate liquidity demand shock realizes, there is a smaller fraction of late consumers, hence less opportunity to exploit the high returns from the productive investment, thus the average utility in the high aggregate liquidity demand state (state H) must be lower than in the low aggregate liquidity demand state (state L).

Thus, $\frac{\partial V_{ND}^{GB^F}}{\partial p}, \frac{\partial V_{D}^{GB^F}}{\partial p} \leq 0$

Second, at any intersection point in p, that is: $\check{p}$ such that $V_D^{GB^F}(\check{p}) \equiv V_{ND}^{GB^F}(\check{p})$, I obtain:

(i) $u(c_H^{FL}) < \left[(\gamma+\alpha)u(c_{1,ND})+(1-\gamma-\alpha)u(c_{2H,ND})\right]$, since partial liquidation must yield higher utility than default in state H, and

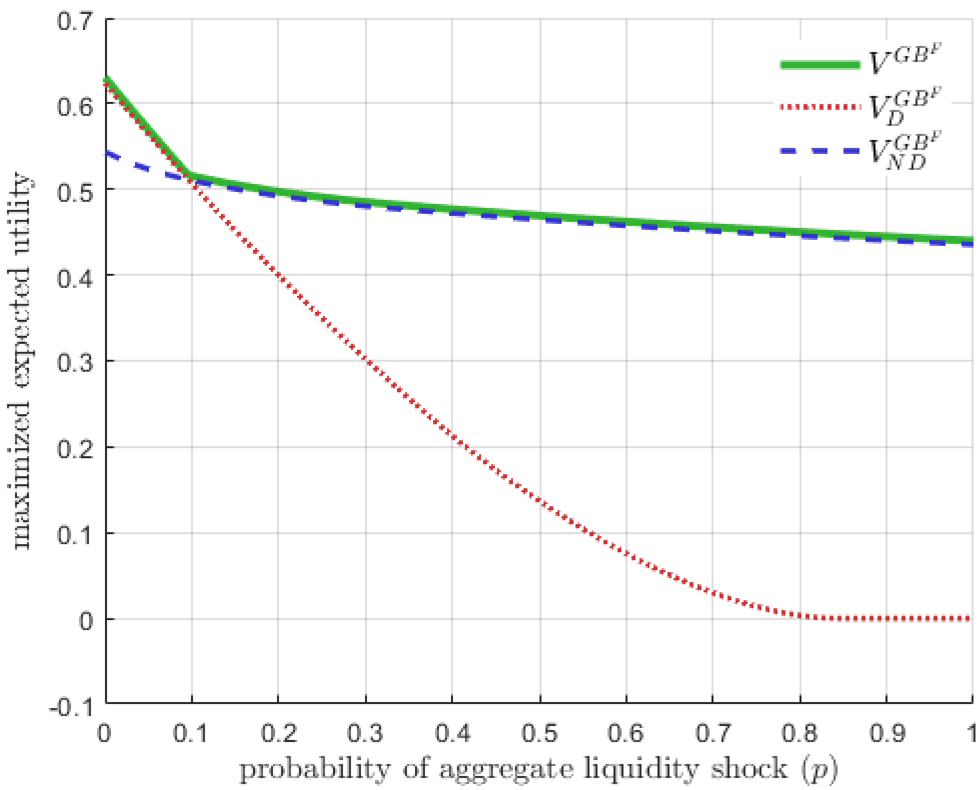
(ii) $\left[\gamma u(c_{1L,D})+(1-\gamma)u(c_{2L,D})\right] \geq \left[\gamma u(c_{1,ND})+(1-\gamma)u(c_{2L,ND})\right]$, since in state L the FL allocation is less constrained than the NFL allocation, which is constrained by facing a trade-off with the allocation in state H.

Thus I conclude that $\frac{\partial V_{D}^{GB^F}}{\partial p} < \frac{\partial V_{ND}^{GB^F}}{\partial p}$ at any $\check{p}$ where $V_D^{GB^F}(\check{p}) = V_{ND}^{GB^F}(\check{p})$. Therefore, the intersection point $\check{p}$ is unique. Figure A.1 illustrates this result.

Combining the results above: there always exists a $\check{p}$ (which may be equal to zero) such that $V_D^{GB^F}(p) \geq V_{ND}^{GB^F}(p) \forall p \leq \check{p}$ and $V_D^{GB^F}(p) < V_{ND}^{GB^F}(p) \forall p > \check{p}$. In words, there always exists a probability of the aggregate liquidity demand shock, $\check{p} \geq 0$ whereby the default regime of the fully informed global bank is superior for all $p \leq \check{p}$, but the no-default regime is superior for all larger p. Since the switch from one regime to another occurs at a unique intersection point of the distinct value functions of the regimes that make up the complete problem, the value function that defines the solution to the complete problem is absolutely continuous, even though the optimal choice variables and consumption allocations may jump discontinuously at the transition from one optimal regime to the other.

Finally, I characterize $\check{p}$ as a function of α, r, and γ. At $\alpha = 0$, for any $\gamma \in [0,1]$, $r \leq 1$ and $R > 1$, the no-default regime must be weakly better than the default regime. This is because the consumption allocation in the default regime in state H is bounded above by 1, while the

Figure A.1: The value function of the complete problem of the global bank with observable types. The figure shows the individual value functions for each regime over the full range of the probability of the aggregate liquidity demand shock p, as well as the value function of the complete problem which is the upper envelope of the two regime specific value functions. When the probability is low ($p < 0.10$) the default regime is superior (with value function $V_D^{GB^F}$, the dotted red line). When the probability is high ($p > 0.10$) the no-default regime is superior (with value function $V_{ND}^{GB^F}$, the dashed blow line). The value function of the complete problem, V^{GB^F} (solid green line), is the absolutely continuous upper envelope of the two regime specific value functions. Parameters: $u(c) = \frac{c^{1-\rho}-1}{1-\rho}$, $\rho = 2$, $R = 5$, $r = 0.1$, $\gamma = 0.5$ and $\varepsilon = 0.1$.



state H allocation in the no-default regime is effectively no more constrained than in L. In the (near) zero aggregate risk case, it must be that $c_{2L} \geq c_{2H} > c_1$ otherwise the investment is not optimally exploited, because (almost) any allocation that satisfies $R > c_{2L} = c_{2H} \geq c_1 \geq r$ is feasible and weakly better than $R > c_{2L} \geq c_{2H} = c_1 = r$. Hence, there must exist a non-zero measure neighbourhood around $\alpha = 0$, where default regime is not optimal. Similarly, when $\alpha = 1 - \gamma$ (or in a neighbourhood), there are (almost) no late consumers to provide for in H, hence the shadow cost of maintaining constant early consumption while providing for some $c_{2H} \neq c_1$ is very high. As such, the default regime is superior in this region (for p low enough). Similar arguments hold for γ. Thus I can define: $\breve{\alpha} \equiv \left\{\alpha | V_{ND}^{GB^F}(\alpha) = V_D^{GB^F}(\alpha)\right\}$, with $V_{ND}^{GB^F}(\alpha) < V_D^{GB^F}(\alpha)$ for $\alpha > \breve{\alpha}$. Similarly, $\breve{\gamma} \equiv \left\{\gamma | V_{ND}^{GB^F}(\gamma) = V_D^{GB^F}(\gamma)\right\}$, with $V_{ND}^{GB^F}(\gamma) < V_D^{GB^F}(\gamma)$ for $\gamma > \breve{\gamma}$.

At $r = 1$ (or in a neighbourhood), partial liquidation of the investment is (almost) without penalty relative to liquidity in financing early consumption. Hence $c_{2L}, c_{2H} > c_1$ can be supported in the no-default regime, which thus dominates the default regime. On the other hand, when r is in a neighbourhood of 0, maintaining the non-contingent early payout in the no-default regime is very costly, and thus the default regime dominates (for low enough p and high enough α and γ). I can thus define $\breve{r} \equiv \left\{r | V_{ND}^{GB^F}(r) = V_D^{GB^F}(r)\right\}$, with $V_{ND}^{GB^F}(r) > V_D^{GB^F}(r)$ for $r > \breve{r}$.

A.4 Proof of Proposition 4

First, suppose that the parameters are such that default in state H is optimal for a fully informed global bank. The allocation in state H is trivially incentive compatible, since $c^*_{2H} = c^*_{1H} = c^*_H$ by construction (pro-rata resolution under default). The incentive compatibility of the allocation in L: $c^*_{2L} \geq c^*_{1L}$, follows directly from the characterization in the proof of Proposition 2:

(i) If $p \geq \overline{p}_D^{GB^F}$, $c^*_{2L} = c^*_{1L} = 1$;

(ii) If $\underline{p}_D^{GB^F} \leq p < \overline{p}_D^{GB^F}$, $c^*_{2L} = c^*_{1L} > 1$;

(iii) If $p < \underline{p}_D^{GB^F}$, $c^*_{2L} > c^*_{1L}$.

Second, suppose that the parameters are such that default in state H is not optimal for a fully informed global bank (see Proposition 1). I consider the three cases characterized in Proposition 1 in turn.

Case 1: $p \geq \overline{p}_{ND}^{GB^F}$ **so that** $\lambda^* = 0$ **and** $e^* > 0$**.** Proposition 1 establishes that the optimal level of excess liquidity e^* in this case satisfies $e^* < \tilde{e}$, where at $\tilde{e}$, $c_1(\tilde{e}) = c_{2L}(\tilde{e}) = c_{2H}(\tilde{e})$. By construction, $\frac{\partial c_1}{\partial e} > 0$, $\frac{\partial c_{2L}}{\partial e} < 0$ and $\frac{\partial c_{2H}}{\partial e} < 0$, therefore $e^* < \tilde{e}$ implies $c_{2L}^*, c_{2H}^* > c_1^*$ in this case. Hence, $p \geq \overline{p}_{ND}^{GB^F}$ is sufficient for incentive compatibility.

Case 2: $\underline{p}_{ND}^{GB^F} < p < \overline{p}_{ND}^{GB^F}$ **so that** $e^* > 0$ **and** $\lambda^* > 0$**.** Conditions (A.5) and (A.6) holding with equality imply:

$$u'\left(c_1^*\right) = \left(1-p\right) \frac{\left(\alpha\left(R-1\right) + \gamma\left(R-r\right)\right)}{\left(1-r\right)\left(\gamma + p\alpha\right)} u'\left(c_{2L}^*\right) \tag{A.18}$$

and

$$u'\left(c_1^*\right) = p \frac{R}{r} \frac{\left(\alpha\left(R-1\right) + \gamma\left(R-r\right)\right)}{\left(R-1\right)\left(\gamma + p\alpha\right)} u'\left(c_{2H}^*\right) \tag{A.19}$$

Condition (A.18) implies $c_{2L}^* \geq c_1^*$ if $R \geq \frac{1-pr}{1-p}$. Since $\frac{1-pr}{1-p} \leq 1$ and $R \geq 1$ this condition is always satisfied. Therefore, $c_{2L}^* \geq c_1^*$ in this case.

Condition (A.19) implies that $c_{2H}^* < c_1^*$ in this case only if

$$p < \hat{p}_{IC}^{(1)} \equiv \frac{\gamma r (R-1)}{(R-r)(\alpha(R-1) + \gamma R)}. \tag{A.20}$$

Hence, $p \geq \hat{p}_{IC}^{(1)}$ is necessary and sufficient for incentive compatibility in this case.

Case 3: $p < \underline{p}_{ND}^{GB^F}$ **so that** $\lambda^* > 0$ **and** $e^* = 0$**.** Proposition 1 establishes that $c_{2L}^* > c_{2H}^*$ in this case. Combined with $c_1^* < \max\left\{c_{2L}^*, c_{2H}^*\right\}$, this implies $c_{2L}^* > c_1^*$. Given that $e^* = 0$, the accounting identities imply that $c_{2H}^* < c_1^*$ if

$$\lambda^* > \hat{\lambda} \equiv \frac{\alpha R}{r(\gamma R + 1 - \alpha - \gamma) + \alpha R}.$$

Since Proposition 1 establishes that $\frac{\partial \lambda^*}{\partial p} < 0$ in this case, and by construction $\frac{\partial c_1}{\partial \lambda} > 0$ and $\frac{\partial c_{2H}}{\partial \lambda} < 0$, there exists a unique $\hat{p}_{IC}^{(2)} \equiv \{p \,|\, \lambda^*(p) = \hat{\lambda}\}$ such that the consumption allocation is incentive compatible if and only if $p \geq \hat{p}_{IC}^{(2)}$.

Last, I establish continuity of the incentive compatibility boundary across different cases: i.e. I show $\lim_{p \nearrow \underline{p}_{ND}^{GB^F}} \hat{p}_{IC}^{(2)} = \lim_{p \searrow \underline{p}_{ND}^{GB^F}} \hat{p}_{IC}^{(1)}$. Suppose by contradiction that $\lim_{p \nearrow \underline{p}_{ND}^{GB^F}} \hat{p}_{IC}^{(2)} > \lim_{p \searrow \underline{p}_{ND}^{GB^F}} \hat{p}_{IC}^{(1)}$. Then there must exist a sequence $p'_n < \hat{p}_{IC}^{(2)}$ converging to $\dot{p} = \underline{p}_{ND}^{GB^F} < \hat{p}_{IC}^{(2)}$ from below, so that $c_{2H}^*(p'_n) < c_1^*(p'_n)$ for all p'_n. There must also be a sequence $p''_n > \hat{p}_{IC}^{(1)}$ converging to $\dot{p} = \underline{p}_{ND}^{GB^F} > \hat{p}_{IC}^{(1)}$ from above, so that $c_{2H}^*(p''_n) > c_1^*(p''_n)$ for all p''_n. But then there must be a discontinuity in the consumption allocation (and hence in the optimal portfolio) at $\dot{p}$, which contradicts the result from the theorem of the maximum that the maximizers of a continuous problem must be continuous. A similar argument would hold for the opposite inequality.

Thus, I conclude that there is a unique, continuous boundary $\hat{p}_{IC}$ where the optimal allocation is characterized by $c_{2H}^*(\hat{p}_{IC}) = c_1^*(\hat{p}_{IC})$, where

$$\hat{p}_{IC} = \begin{cases} \hat{p}_{IC}^{(1)} & \text{if } p > \underline{p}_{ND}^{GB^F} \\ \hat{p}_{IC}^{(2)} & \text{if } p \leq \underline{p}_{ND}^{GB^F} \end{cases}$$

Figure 3.5 illustrates the regions where the hierarchy of consumption levels differs, as well as the piece-wise defined but continuous nature of the incentive compatibility bound, $\hat{p}_{IC}$.

A.5 Proof of Proposition 5

The allocation chosen by a global bank that cannot observe consumer types is identical to the allocation chosen by a global bank that can observe consumer types whenever the latter is incentive compatible. Proposition 4 establishes that the optimal allocation chosen by a global bank that can observe consumer types is always incentive compatible when the default regime is optimal (i.e. when $p \leq \breve{p}$). Therefore I need to consider only the no-default regime.

The no-default regime of a global bank that cannot observe consumer types (called a global bank under asymmetric information, denoted by GB^A), is subject to constraint that early consumption must be non-state-contingent $c_{1L} = c_{1H} = c_1$. The problem in this regime (P2a) is defined by the value function:

$$V_{ND}^{GB^A} \equiv \max_{\{c_1, c_{2L}, c_{2H}\}} (1-p)\Big[\gamma u(c_1) + (1-\gamma) u(c_{2L})\Big] + p\Big[(\gamma+\alpha) u(c_1) + (1-\gamma-\alpha) u(c_{2H})\Big] \tag{P2a}$$

$$\text{subject to:} \qquad \begin{aligned} & x + y = 1, \\ & c_{2L} = \frac{y - \gamma c_1 + Rx}{1-\gamma}, \\ & c_{2H} = \frac{R}{1-\gamma-\alpha}\left(x - \frac{(\gamma+\alpha)c_1 - y}{r}\right), \\ & c_{2L}, c_{2H} \geq c_1. \end{aligned}$$

As in the case of a global bank that can observe consumer types, it is convenient to state the problem in terms of excess liquidity in state L and partial liquidation in state H. In principle, the optimal allocation of a global bank that cannot observe consumer types will have the same three cases as that of a global bank that can observe consumer types. Therefore, there are corner solutions: for some parameter ranges, either of the choice variables may be zero at the optimum. However, in the global bank problem under full information, Proposition 4 shows that there is never a violation of incentive compatibility when zero partial liquidation is optimal (i.e. when $p \geq \overline{p}_{ND}^{GB^F}$). Therefore I need to consider only the other two cases ($p \leq \underline{p}_{ND}^{GB^F}$ and $\underline{p}_{ND}^{GB^F} < p < \overline{p}_{ND}^{GB^F}$).

Proposition 4 shows that incentive compatibility is violated in these cases if and only if $p < \hat{p}_{IC}$. Moreover, incentive compatibility is violated only in the state H allocation. Consider the parameter values where this is the case. When incentive compatibility is violated in state H in the allocation of a global bank under full information (i.e. when $c_{2H} < c_1$), the closest weakly incentive-compatible allocation to that that a global bank can achieve under asymmetric information is characterized by $c_{2H} = c_1 \equiv c_H$, in which case the expected utility reduces to $W_{ND}^{GB^A} = \left(p + \left(1-p\right)\gamma\right) u\left(c_H\right) + \left(1-p\right)\left(1-\gamma\right) u\left(c_{2L}\right)$. The character of the solution of the global bank problem under asymmetric information differs between the two cases identified by a bound on the probability of the aggregate liquidity demand shock, denoted $\underline{p}_{ND}^{GB^A} \in (0, \underline{p}_{ND}^{GB^F})$.

Case 1: $p \leq \underline{p}_{ND}^{GB^A}, \hat{p}_{IC}$**, therefore** $\lambda^* > 0$ **and** $e^* = 0$**.** In this case, there is no default or excess liquidity in L, so that the only way to ensure constant consumption in $t = 1$ is for $\lambda = \frac{\alpha}{r\gamma} y$. Adding the accounting identity $c_1 = \frac{y + r\lambda}{\gamma + \alpha} = \frac{(1-y-\lambda)R}{1-\gamma-\alpha} = c_{2H}$ fully determines the solution, which is independent of the utility function and the probability of the aggregate liquidity shock:

$$
\begin{aligned}
y^* &= \frac{\gamma r R}{\gamma r R + \alpha R + r(1-\gamma-\alpha)}, \\
c_1^* = c_{2H}^* &= \frac{y^*}{\gamma}, \quad c_{2L}^* = \frac{R(1-y^*)}{1-\gamma}.
\end{aligned}
$$

Case 2: $\underline{p}_{ND}^{GB^A} < p < \hat{p}_{IC}$**, so** $e^* > 0$ **and** $\lambda^* > 0$**.** Using $c_1 = c_{2H}$, the problem can be stated in terms of y alone:

$$
\begin{aligned}
e &= \frac{y(r(\gamma R + 1 - \alpha - \gamma) + \alpha R) - \gamma r R}{r(1-\alpha-\gamma) + R(\alpha+\gamma)}, \quad \lambda = \frac{R(\alpha+\gamma) - y((R-1)(\alpha+\gamma)+1)}{r(1-\alpha-\gamma)+R(\alpha+\gamma)} \\
c_H &= \frac{(1-r)Ry + rR}{r(1-\alpha-\gamma)+R(\alpha+\gamma)} \\
c_{2L} &= \frac{rR(1-\alpha-2\gamma) + R^2(\alpha+\gamma)}{(1-\gamma)(r(1-\alpha-\gamma)+R(\alpha+\gamma))} \\
&\quad - \frac{y\left(rR(1-\alpha-2\gamma) + R^2(\alpha+\gamma) - \alpha R - r(1-\alpha-\gamma)\right)}{(1-\gamma)(r(1-\alpha-\gamma)+R(\alpha+\gamma))},
\end{aligned}
$$

yielding the optimality condition $u'(c_H) = \frac{(1-p)(r(R(1-\alpha-2\gamma)-(1-\alpha-\gamma))+R(R(\alpha+\gamma)-\alpha))}{(1-r)R(\gamma+(1-\gamma)p)} u'(c_{2L})$. Since $\frac{\partial c_H}{\partial \lambda} < 0$ and $\frac{\partial c_{2L}}{\partial \lambda} > 0$, there exists a unique $\lambda^* > 0$ that solves this problem.

The proof of the existence and uniqueness of bound $\underline{p}_{ND}^{GB^A}$ under asymmetric information is similar to the proof of the existence and uniqueness of the bound $\underline{p}_{ND}^{GB^F}$ under full information (see section 1) and is therefore omitted.

Since a global bank under asymmetric information must offer $c_{2H} = c_1$ when a global bank under full information offers $c_{2H} < c_1$, a global bank under asymmetric information optimally chooses less partial liquidation than under full information. For the same reasons as in the full information case, when p is low, ex-post inefficiency in H is less costly than in L, thus, a global bank under asymmetric information uses only partial liquidation in a neighbourhood of $p = 0$. However, since a global bank uses less partial liquidation under asymmetric information than under full information, as p increases it becomes efficient to start using some excess liquidity sooner under asymmetric information than under full information; therefore, $\underline{p}_{ND}^{GB^A} < \underline{p}_{ND}^{GB^F}$.

A.6 Proof of proposition 6

The effective probability at $t = 0$ of being an early consumer is $p' \equiv \gamma + p\alpha$. The problem of the consumers is to choose their portfolio to maximize their expected utility in autarky:

$$\max_{x \in [0,1]} p'\, u(c_1(x)) + \left(1 - p'\right) u(c_2(x))$$

If $p'\,(1 - r)\, u'\big(c_1(x)\big) > (1 - p')\,(R - 1)\, u'\big(c_2(x)\big)$ for all x, then it is optimal for the consumer not to invest at all. Thus, $c_1^{Aut} = c_2^{Aut} = 1$ and the optimality condition reduces to an inequality constraint on the effective probability, $p' > \overline{p}^{Aut} \equiv \frac{R-1}{R-r} \in (0, 1)$.

If $p'\,(1-r)\, u'\big(c_1(x)\big) < (1-p')\,(R-1)\, u'\big(c_2(x)\big)$ for all x, it is optimal for the consumer to fully invest. Thus, $c_1^{Aut} = r$ and $c_2^{Aut} = R$ and the optimality condition again reduces to an inequality constraint on the effective probability, $p' < \underline{p}^{Aut} \equiv \frac{(R-1)u'(R)}{(R-1)u'(R)+(1-r)u'(r)} \in (0, 1)$.

Otherwise, there must exist an $x \in (0, 1)$ where $p'\,(1 - r)\, u'\big(c_1(x)\big) = (1 - p')\,(R - 1)\, u'\big(c_2(x)\big)$, which implicitly defines the optimal level of investment $0 < x^{Aut} < 1$. This implies that the consumption levels are $r < c_1^{Aut} < 1 < c_2^{Aut} < R$.

Lastly, $\underline{p}^{Aut} < \overline{p}^{Aut}$ because $\underline{p}^{Aut} \equiv \frac{(R-1)u'(R)}{(R-1)u'(R)+(1-r)u'(r)} < \frac{(R-1)u'(R)}{(1-r)u'(r)} < \frac{(R-1)u'(1)}{(1-r)u'(1)} \equiv \overline{p}^{Aut}$.

Appendix B

Recursively Weighted Least Squares

To test my replication of the Allen and Gale (2000) results, I use recursively weighted least squares (RWLS) regressions with a *fair* weighting scheme as implemented in Matlab (2017).

RWLS proceeds in the following steps for an arbitrary linear regression.

step 1: Initialize estimates by estimating the regression by ordinary least squares, and construct the following statistics from the estimated residuals u_i of the regression:

(i) the vector of leverage values h_i, and

(ii) an estimate of the standard deviation of the error term, given by the bias corrected median absolute deviation (MAD) estimator: $\sigma = \frac{MAD}{0.6745}$,
where $MAD = median(|u_i - median(u_i)|)$.

step 2: Construct a weight for each observation. I use the fair weighting scheme, which assigns weight ω_i to observation i. The weight is defined as:

$$\omega_i = \frac{1}{1 + |\phi_i|},$$

where

$$\phi_i = \frac{u_i}{1.4\sigma\sqrt{1 - h_i}}.$$

The factor 1.4 in the definition of ϕ_i is the default tuning parameter suggested by the algorithm in Matlab (2017). Thus, the weights on observations with larger residuals are smaller than the weights on observations with smaller residuals.

step 3: Perform weighted least squares using the observation weights ω_i and recalculate weights as in step 2.

step 4: Repeat step 3 until convergence (i.e. until the change in weights falls below a critical threshold).